

## ΑΝΑΣΚΟΠΗΣΗ REVIEW

### CONSORT 2010, explanation and elaboration

### Updated guidelines for reporting parallel group randomized trials

### CONSORT 2010, επεξήγηση και επεξεργασία Αναθεωρημένες οδηγίες για την καταγραφή των τυχαιοποιημένων δοκιμών παράλληλων ομάδων

Ελληνική μετάφραση – επιμέλεια:

Σ. Κόκκαλη, Χ. Μπίρτσου, Α.Μ. Χάιδιτς

Εργαστήριο Υγιεινής, Ιατρική Σχολή, Αριστοτέλειο Πανεπιστήμιο  
Θεσσαλονίκης, Θεσσαλονίκη

Ιδιαίτερα πειστικά δεδομένα δείχνουν ότι η ποιότητα καταγραφής των τυχαιοποιημένων κλινικών δοκιμών (ΤΚΔ) με ομάδα ελέγχου δεν είναι η βέλτιστη. Χωρίς διαφανή καταγραφή, οι αναγνώστες δεν μπορούν να κρίνουν την αξιοπιστία και την εγκυρότητα των ευρημάτων των δοκιμών ούτε να εξάγουν τις πληροφορίες για συστηματικές ανασκοπήσεις. Πρόσφατες μεθοδολογικές αναλύσεις δείχνουν ότι η ανεπαρκής καταγραφή και ο μη επαρκής σχεδιασμός συνδέονται με τις εσφαλμένες εκτιμήσεις των θεραπευτικών αποτελεσμάτων. Τέτοιο συστηματικό σφάλμα βλάπτει σοβαρά τις ΤΚΔ, οι οποίες θεωρούνται χρυσός κανόνας για την αξιολόγηση των παρεμβάσεων λόγω της δυνατότητάς τους να ελαχιστοποιούν ή να αποφεύγουν το σφάλμα. Μια ομάδα επιστημόνων και εκδοτών ανέπτυξε τη δήλωση CONSORT (Consolidated Standards of Reporting Trials, Ενισχυμένα Πρότυπα Αναφοράς Δοκιμών) για να βελτιώσει την ποιότητα καταγραφής των ΤΚΔ. Αρχικά, δημοσιεύτηκε το 1996 και αναθεωρήθηκε το 2001. Η δήλωση αποτελείται από έναν κατάλογο και ένα διάγραμμα ροής που μπορούν να χρησιμοποιήσουν οι συγγραφείς για την καταγραφή των αποτελεσμάτων μιας ΤΚΔ. Πολλά κορυφαία ιατρικά περιοδικά και σημαντικές διεθνείς εκδοτικές ομάδες έχουν επιδοκιμάσει τη δήλωση CONSORT. Η δήλωση διευκολύνει την κριτική αξιολόγηση και την ερμηνεία των ΤΚΔ. Κατά τη διάρκεια της αναθεώρησης του CONSORT το 2001, έγινε σαφές ότι η επεξήγηση και η επεξεργασία των αρχών που υπογραμμίζει η δήλωση CONSORT θα βοηθούσαν τους ερευνητές και άλλους να γράψουν ή να αξιολογήσουν τις δημοσιεύσεις των μελετών. Ένα άρθρο επεξήγησης και επεξεργασίας της δήλωσης CONSORT δημοσιεύτηκε το 2001, παράλληλα με την έκδοση της δήλωσης CONSORT του 2001. Μετά από μια συνεδρίαση ειδικών τον Ιανουάριο του 2007, η δήλωση CONSORT έχει αναθεωρηθεί περαιτέρω και δημοσιεύεται ως δήλωση CONSORT 2010. Η εν λόγω αναθεώρηση βελτιώνει τη διατύπωση και τη σαφήνεια του προηγούμενου καταλόγου και ενσωματώνει συστάσεις σχετικές με θέματα που μόνο πρόσφατα έχουν λάβει αναγνώριση, όπως το σφάλμα επιλεκτικής αναφοράς έκβασης (selective out-

ΑΡΧΕΙΑ ΕΛΛΗΝΙΚΗΣ ΙΑΤΡΙΚΗΣ 2011, 28(Συμπλ 1):10-59  
ARCHIVES OF HELLENIC MEDICINE 2011, 28(Suppl 1):10-59

D. Moher,<sup>1</sup>  
S. Hopewell,<sup>2</sup>  
K.F. Schulz,<sup>3</sup>  
V. Montori,<sup>4</sup>  
P.C. Gøtzsche,<sup>5</sup>  
P.J. Devereaux,<sup>6</sup>  
D. Elbourne,<sup>7</sup>  
M. Egger,<sup>8</sup>  
D.G. Altman<sup>2</sup>

<sup>1</sup>Ottawa Methods Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa Hospital, Ottawa, Ontario, Canada, K1H 8L6

<sup>2</sup>Centre for Statistics in Medicine, University of Oxford, Wolfson College, Oxford, UK

<sup>3</sup>Family Health International, Research Triangle Park, NC 27709, USA

<sup>4</sup>UK Knowledge and Encounter Research Unit, Mayo Clinic, Rochester, MN, USA

<sup>5</sup>The Nordic Cochrane Centre, Rigshospitalet, Blegdamsvej 9, Copenhagen, Denmark

<sup>6</sup>McMaster University Health Sciences Centre, Hamilton, Canada

<sup>7</sup>Medical Statistics Unit, London School of Hygiene and Tropical Medicine, London, UK

<sup>8</sup>Institute of Social and Preventive Medicine (ISPM), University of Bern, Switzerland

CONSORT 2010, explanation and elaboration: Updated guidelines for reporting parallel group randomised trials?

Abstract at the end of the article

#### Λέξεις ευρετηρίου

Διαφάνεια  
Δήλωση CONSORT  
Κριτική αξιολόγηση  
Μεροληψία  
Ποιότητα  
Συστηματικό σφάλμα  
Τυχαιοποιημένη κλινική δοκιμή (ΤΚΔ)

come reporting bias). Αυτό το έγγραφο επεξήγησης και επεξεργασίας –που επιδιώκει να ενισχύσει τη χρήση, την κατανόηση και τη διάδοση της δήλωσης του CONSORT– έχει επίσης αναθεωρηθεί εκτενώς. Παρουσιάζει το νόημα και τη λογική για κάθε νέο και ενημερωμένο στοιχείο του καταλόγου, παρέχοντας παραδείγματα καλής καταγραφής και, όπου είναι δυνατόν, αναφορές σε σχετικές εμπειρικές μελέτες. Συμπεριλαμβάνονται διάφορα παραδείγματα διαγραμμάτων ροής. Η δήλωση CONSORT 2010, το παρόν αναθεωρημένο έγγραφο επεξήγησης και επεξεργασίας, καθώς και η σχετική ιστοσελίδα ([www.consort-statement.org](http://www.consort-statement.org)) αποτελούν χρήσιμες πηγές για τη βελτίωση της καταγραφής των τυχαιοποιημένων δοκιμών.

## 1. ΕΙΣΑΓΩΓΗ

«Όλη η Ιατρική εξαρτάται από τη διαφανή καταγραφή των κλινικών δοκιμών».<sup>1</sup>

Οι καλά σχεδιασμένες και κατάλληλα εκτελεσμένες τυχαιοποιημένες κλινικές δοκιμές (ΤΚΔ) με ομάδα ελέγχου παρέχουν τα πλέον αξιόπιστα δεδομένα για την αποτελεσματικότητα των παρεμβάσεων στη φροντίδα υγείας, ενώ δοκιμές με ανεπαρκή μεθοδολογία συνδέονται με σφάλμα, ειδικά την υπερεκτίμηση των θεραπευτικών αποτελεσμάτων.<sup>2-5</sup> Εσφαλμένα αποτελέσματα από κακοσχεδιασμένες και ανεπαρκώς αναφερόμενες δοκιμές μπορούν να παραπλανήσουν την απόφαση στη φροντίδα υγείας σε όλα τα επίπεδα, από τις αποφάσεις θεραπείας για έναν ασθενή μέχρι τη διαμόρφωση των εθνικών πολιτικών δημόσιας υγείας.

Η κριτική αξιολόγηση της ποιότητας των κλινικών δοκιμών είναι δυνατή μόνο εάν ο σχεδιασμός, η διεξαγωγή και η ανάλυση των ΤΚΔ περιγράφονται επιμελώς και επακριβώς στη δημοσίευση. Μακριά από την ύπαρξη διαφάνειας, η έκθεση των ΤΚΔ είναι συχνά ελλιπής,<sup>6-9</sup> συνδυάζοντας προβλήματα που προκύπτουν από την ανεπαρκή μεθοδολογία.<sup>10-15</sup>

## 2. ΕΛΛΙΠΗΣ ΚΑΙ ΑΝΑΚΡΙΒΗΣ ΚΑΤΑΓΡΑΦΗ

Πολλές ανασκοπήσεις έχουν τεκμηριώσει ανεπάρκειες στις δημοσιεύσεις των κλινικών δοκιμών. Για παράδειγμα, πληροφορίες για τη μέθοδο που χρησιμοποιήθηκε σε μια δοκιμή προκειμένου να ταξινομηθούν οι συμμετέχοντες στις ομάδες σύγκρισης αναφέρθηκαν μόνο στο 21% των 519 δημοσιευμένων δοκιμών που καταχωρήθηκαν στο PubMed το 2000<sup>16</sup> και μόνο στο 34% από τις 616 εργασίες που καταχωρήθηκαν το 2006.<sup>17</sup> Ομοίως, μόνο στο 45% των δημοσιευμένων δοκιμών που καταχωρήθηκαν στο PubMed το 2000 και στο 53% το 2006 καθοριζόταν ένα πρωτεύον καταληκτικό σημείο, και μόνο στο 27% το 2000 και στο 45% το 2006 αναφερόταν ο υπολογισμός μεγέθους του δείγματος. Η καταγραφή είναι όχι μόνο συχνά ελλιπής αλλά και μερικές φορές ανακριβής. Από τις 119 δημοσιεύσεις που δηλώνουν ότι όλοι οι συμμετέχοντες περιελήφθησαν στην ανάλυση στις ομάδες που ταξινομήθηκαν αρχικά (ανάλυση πρόθεσης για θεραπεία, intention-to-treat analysis), στις

15 (13%) αποκλείστηκαν ασθενείς ή δεν αναλύθηκαν όλοι οι ασθενείς, όπως είχαν διανεμηθεί.<sup>18</sup> Πολλές άλλες ανασκοπήσεις έχουν διαπιστώσει ότι η ανεπαρκής καταγραφή είναι συχνή σε περιοδικά ειδικότητας<sup>16-19</sup> και σε περιοδικά που δημοσιεύονται σε γλώσσες εκτός από την αγγλική.<sup>20,21</sup>

Η κατάλληλη τυχαιοποίηση μειώνει το σφάλμα επιλογής κατά την είσοδο στη δοκιμή και είναι το κρίσιμο συστατικό των υψηλής ποιότητας ΤΚΔ.<sup>22</sup> Η επιτυχημένη τυχαιοποίηση εξαρτάται από δύο βήματα: την παραγωγή μιας μη προβλέψιμης ακολουθίας διανομής και την απόκρυψη αυτής της ακολουθίας από τους ερευνητές που εγγράφουν τους συμμετέχοντες (βλέπε πλαίσιο 1).<sup>2,23</sup>

Δυστυχώς, παρά τον κεντρικό ρόλο, η καταγραφή των μεθόδων που χρησιμοποιούνται για τη διανομή των συμμετεχόντων στις παρεμβάσεις, γενικά, είναι επίσης ανεπαρκής. Για παράδειγμα, το 5% των 206 δημοσιεύσεων των υποτιθέμενων ΤΚΔ στα περιοδικά Μαιευτικής και Γυναικολογίας περιέγραφε μελέτες που δεν ήταν αληθινά τυχαιοποιημένες.<sup>23</sup> Η συγκεκριμένη εκτίμηση είναι συντηρητική, καθώς οι περισσότερες δημοσιεύσεις δεν παρέχουν αυτή τη στιγμή επαρκείς πληροφορίες για τη μέθοδο της διανομής.<sup>20,23,30-33</sup>

## 3. ΒΕΛΤΙΩΝΟΝΤΑΣ ΤΗΝ ΥΠΟΒΟΛΗ ΔΗΜΟΣΙΕΥΣΕΩΝ ΤΩΝ ΤΥΧΑΙΟΠΟΙΗΜΕΝΩΝ ΚΛΙΝΙΚΩΝ ΔΟΚΙΜΩΝ: Η ΔΗΛΩΣΗ CONSORT

Οι DerSimonian et al πρότειναν ότι «οι εκδότες θα μπορούσαν να βελτιώσουν πολύ τη δημοσίευση των κλινικών δοκιμών παρέχοντας στους συγγραφείς έναν κατάλογο στοιχείων, στον οποίο προσδοκούσαν να αναφερθούν αυστηρά».<sup>34</sup> Στις αρχές του 1990 δύο ομάδες εκδοτών περιοδικών, διεξαγόντων δοκιμές, και μεθοδολόγων δημοσίευσαν ανεξάρτητα συστάσεις σχετικά με τη δημοσίευση των δοκιμών.<sup>35,36</sup> Σε ένα επόμενο συνοπτικό άρθρο, ο Rennie πίεσε τις δύο ομάδες για να συναντηθούν και να αναπτύξουν ένα κοινό σύνολο συστάσεων.<sup>37</sup> Το αποτέλεσμα ήταν η δήλωση CONSORT (Consolidated Standards of Reporting Trials, Ενισχυμένα Πρότυπα Αναφοράς Δοκιμών).<sup>38</sup>

Η δήλωση CONSORT –ή απλά CONSORT– περιλαμβάνει έναν κατάλογο βασικών στοιχείων που πρέπει να περιληφθούν στις δημοσιεύσεις των ΤΚΔ και ένα διάγραμμα για

την τεκμηρίωση της ροής των συμμετεχόντων μέσα από μια δοκιμή. Είναι στοχευμένο στις αρχικές δημοσιεύσεις των ΤΚΔ με δύο ομάδες, παράλληλου σχεδιασμού. Το μεγαλύτερο τμήμα του CONSORT σχετίζεται επίσης με μια ευρύτερη κατηγορία σχεδιασμών μελετών, όπως μη υπεροχής, ισοδυναμίας, παραγοντικές, συρροής και διασταυρούμενες δοκιμές. Έχουν δημοσιευτεί επεκτάσεις του καταλόγου του CONSORT για την καταγραφή των δοκιμών με μερικούς από αυτούς τους σχεδιασμούς,<sup>39-41</sup> καθώς έχουν δημοσιευτεί και επεκτάσεις για την καταγραφή συγκεκριμένων τύπων δεδομένων (βλαβερές συνέπειες<sup>42</sup>), ειδών παρεμβάσεων (μη φαρμακολογικές θεραπείες,<sup>43</sup> παρεμβάσεις με βότανα<sup>44</sup>) και περιλήψεων.<sup>45</sup>

Ο στόχος του CONSORT είναι να παράσχει καθοδήγηση στους συγγραφείς για το πώς να βελτιώσουν την καταγραφή των δοκιμών τους. Οι δημοσιεύσεις των δοκιμών χρειάζεται να είναι σαφείς, πλήρεις και διαφανείς. Οι αναγνώστες, οι κριτές (peer reviewers) και οι εκδότες μπορούν επίσης να χρησιμοποιήσουν το CONSORT για να τους βοηθήσει στην κριτική αξιολόγηση και στην ερμηνεία των δημοσιεύσεων των ΤΚΔ. Εν τούτοις, το CONSORT δεν προορίστηκε να χρησιμοποιηθεί ως όργανο ποιοτικής αξιολόγησης. Αντίθετα, το περιεχόμενο του CONSORT εστιάζει στα στοιχεία που σχετίζονται με την εσωτερική και την εξωτερική εγκυρότητα των δοκιμών. Πολλά στοιχεία που δεν αναφέρονται διεξοδικά στο CONSORT πρέπει επίσης να περιληφθούν σε μια αναφορά, όπως οι πληροφορίες για την έγκριση από μια επιτροπή βιοηθικής, η λήψη της μετά

από ενημέρωση συγκατάθεσης από τους συμμετέχοντες και, στις σχετικές περιπτώσεις, η ύπαρξη μιας επιτροπής ασφάλειας και επιτήρησης δεδομένων (data safety and monitoring committee). Επί πλέον, οποιεσδήποτε άλλες πτυχές μιας δοκιμής που αναφέρονται πρέπει να καταγραφούν κατάλληλα, όπως οι πληροφορίες σχετικά με την ανάλυση κόστους-αποτελεσματικότητας.<sup>46-48</sup>

Από τη δημοσίευσή του το 1996, το CONSORT έχει υποστηριχθεί από περισσότερα από 400 περιοδικά ([www.consort-statement.org](http://www.consort-statement.org)) και διάφορες εκδοτικές ομάδες, όπως η Διεθνής Επιτροπή Εκδοτών Ιατρικών Περιοδικών (International Committee of Medical Journal Editors).<sup>49</sup> Η εισαγωγή του CONSORT μέσα στα περιοδικά συνδέεται με τη βελτιωμένη ποιότητα των δημοσιευμένων ΤΚΔ.<sup>17,50,51</sup> Εν τούτοις, το CONSORT είναι μια εξελισσόμενη πρωτοβουλία και η δήλωσή του αναθεωρείται περιοδικά.<sup>3</sup> Το CONSORT αναθεωρήθηκε τελευταία πριν από 9 χρόνια, το 2001.<sup>52-54</sup> Από τότε η βάση των δεδομένων ενημέρωσης του CONSORT έχει αυξηθεί σημαντικά. Τα εμπειρικά δεδομένα έχουν δώσει έμφαση στις νέες ανησυχίες σχετικά με την καταγραφή των ΤΚΔ, όπως η επιλεκτική καταγραφή της έκβασης.<sup>55-57</sup> Μία συνεδρίαση της ομάδας του CONSORT επομένως συγκλήθηκε τον Ιανουάριο του 2007, στον Καναδά, για να αναθεωρήσει τη δήλωση CONSORT του 2001, καθώς και το συνοδευτικό έγγραφο εξήγησης και επεξεργασίας. Ο αναθεωρημένος κατάλογος παρουσιάζεται στον πίνακα 1 και το διάγραμμα ροής, μη αναθεωρημένο, στην εικόνα 1.<sup>52-54</sup>

**Πίνακας 1.** Κατάλογος στοιχείων που θα πρέπει να περιληφθούν στην αναφορά μιας τυχαίοποιημένης δοκιμής, σύμφωνα με τη δήλωση CONSORT 2010.\*

#### Τίτλος και περίληψη

- 1α Αναγνώριση της δοκιμής ως τυχαίοποιημένη στον τίτλο
- 1β Δομημένη σύνοψη του σχεδιασμού, των μεθόδων, των αποτελεσμάτων και των συμπερασμάτων της δοκιμής (για ειδική καθοδήγηση, ανατρέξτε στο CONSORT για περιλήψεις)

#### Εισαγωγή

- Υπόβαθρο και στόχοι
  - 2α Επιστημονικό υπόβαθρο και η εξήγηση της λογικής
  - 2β Ειδικοί στόχοι ή υποθέσεις

#### Μέθοδοι

- Σχεδιασμός δοκιμής
  - 3α Περιγραφή του σχεδιασμού της δοκιμής (όπως παράλληλος, παραγοντικός), περιλαμβανομένης της αναλογίας διανομής
  - 3β Σημαντικές αλλαγές στις μεθόδους ύστερα από την έναρξη της δοκιμής (όπως στα κριτήρια καταλληλότητας), με αιτιολόγηση
- Συμμετέχοντες
  - 4α Κριτήρια καταλληλότητας για τους συμμετέχοντες
  - 4β Συνθήκες και τόποι στους οποίους συλλέχθηκαν τα δεδομένα
- Παρεμβάσεις
  - 5 Οι παρεμβάσεις της κάθε ομάδας με αρκετή λεπτομέρεια, έτσι ώστε να είναι εφικτή η επανάληψη, περιλαμβανομένου του πώς και πότε πραγματικά χορηγήθηκαν
- Εκβάσεις
  - 6α Ολοκληρωμένος προσδιορισμός των προκαθορισμένων κύριων και δευτερευόντων μέτρων έκβασης, περιλαμβανομένου του πώς και πότε αξιολογήθηκαν
  - 6β Οποιοσδήποτε αλλαγές στις εκβάσεις της δοκιμής ύστερα από την έναρξη της δοκιμής, με αιτίες

\*Προτείνουμε έντονα να διαβάζεται αυτή η δήλωση σε συνδυασμό με το CONSORT 2010, Εξήγηση και Επεξεργασία (CONSORT 2010 Explanation and Elaboration) για σημαντικές διευκρινίσεις σε όλα τα στοιχεία. Επίσης, αν είναι σχετικό, προτείνεται η ανάγνωση των επεκτάσεων του CONSORT για τυχαίοποιημένες δοκιμές συρροής,<sup>40</sup> δοκιμές μη υπεροχής και ισοδυναμίας,<sup>39</sup> μη φαρμακολογικές θεραπείες,<sup>43</sup> βοτανικές παρεμβάσεις<sup>44</sup> και πραγματικές δοκιμές (pragmatic trials).<sup>41</sup> Πρόσθετες επεκτάσεις θα είναι διαθέσιμες: Γι' αυτές και για άλλες επίκαιρες αναφορές σχετικές με αυτόν τον κατάλογο, ανατρέξτε στο δικτυακό τόπο [www.consort-statement.org](http://www.consort-statement.org)

**Πίνακας 1.** (συνέχεια) Κατάλογος στοιχείων που θα πρέπει να περιληφθούν στην αναφορά μιας τυχαιοποιημένης δοκιμής, σύμφωνα με τη δήλωση CONSORT 2010.\*

Μέγεθος δείγματος	7α	Πώς καθορίστηκε το μέγεθος του δείγματος
	7β	Όπου είναι εφαρμόσιμο, επεξήγηση των ενδιάμεσων αναλύσεων και οδηγιών διακοπής
<b>Τυχαιοποίηση</b>		
Δημιουργία ακολουθίας	8α	Μέθοδος που χρησιμοποιήθηκε για τη δημιουργία της ακολουθίας τυχαίας διανομής
	8β	Τύπος τυχαιοποίησης <sup>37</sup> λεπτομέρειες τυχόν περιορισμών (όπως χρήση μπλοκ [blocking] και μέγεθος μπλοκ)
Μηχανισμός απόκρυψης διανομής	9	Μηχανισμός που χρησιμοποιήθηκε για την εφαρμογή της ακολουθίας τυχαίας διανομής (όπως διαδοχικά αριθμημένοι περιέκτες), περιγράφοντας οποιαδήποτε βήματα που τηρήθηκαν για την απόκρυψη της ακολουθίας έως την εφαρμογή των παρεμβάσεων
Εφαρμογή	10	Ποιος δημιούργησε την ακολουθία τυχαίας διανομής, ποιος ενέγραψε τους συμμετέχοντες και ποιος ταξινόμησε τους συμμετέχοντες στις παρεμβάσεις
Τυφλοποίηση	11α	Αν έγινε, ποιος ήταν τυφλοποιημένος μετά την ταξινόμηση των παρεμβάσεων (για παράδειγμα, οι συμμετέχοντες, οι προμηθευτές φροντίδας, αυτοί που αξιολογούν τις εκβάσεις) και πώς εφαρμόστηκε
	11β	Αν κρίνεται σχετικό, περιγραφή της ομοιότητας των παρεμβάσεων
Στατιστικές μέθοδοι	12α	Στατιστικές μέθοδοι που εφαρμόστηκαν για τη σύγκριση των ομάδων, για τις κύριες και τις δευτερεύουσες εκβάσεις
	12β	Μέθοδοι για πρόσθετες αναλύσεις, όπως αναλύσεις υποομάδων και σταθμισμένες αναλύσεις
<b>Αποτελέσματα</b>		
Ροή συμμετεχόντων (ένα διάγραμμα συστήνεται έντονα)	13α	Για κάθε ομάδα, ο αριθμός των συμμετεχόντων, που ταξινομήθηκαν τυχαία, που έλαβαν την προορισμένη παρέμβαση και αναλύθηκαν για την κύρια έκβαση
	13β	Για κάθε ομάδα, απώλειες και αποκλεισμοί μετά από την τυχαιοποίηση, μαζί με αιτίες
Στρατολόγηση	14α	Ημερομηνίες που καθορίζουν τις περιόδους στρατολόγησης και παρακολούθησης
	14β	Γιατί η δοκιμή τερματίστηκε ή διακόπηκε
Βασικά αρχικά δεδομένα	15	Ένας πίνακας στον οποίο να φαίνονται τα βασικά δημογραφικά και κλινικά χαρακτηριστικά για κάθε ομάδα
Αριθμοί που αναλύθηκαν	16	Για κάθε ομάδα, ο αριθμός των συμμετεχόντων (παρονομαστής) που συμπεριλήφθηκαν σε κάθε ανάλυση και κατά πόσο η ανάλυση έλαβε χώρα στις αρχικά ταξινομημένες ομάδες.
Εκβάσεις και εκτίμηση	17α	Για κάθε κύρια και δευτερεύουσα έκβαση, αποτελέσματα για κάθε ομάδα, καθώς και το εκτιμώμενο μέγεθος επίδρασης και η ακρίβειά του (όπως τα 95% διαστήματα εμπιστοσύνης)
	17β	Για διχότομες εκβάσεις, συστήνεται η παρουσίαση τόσο των απόλυτων, όσο και των σχετικών μεγεθών επίδρασης
Συμπληρωματικές αναλύσεις	18	Αποτελέσματα κάθε άλλης ανάλυσης, περιλαμβανομένης της ανάλυσης υποομάδων και σταθμισμένης ανάλυσης, διακρίνοντας την προκαθορισμένη από τη διερευνητική
Βλαβερές συνέπειες	19	Όλες οι σημαντικές βλαβερές συνέπειες ή οι ακούσιες επιδράσεις σε κάθε ομάδα (για εξειδικευμένη καθοδήγηση, ανατρέξτε στο CONSORT για βλαβερές συνέπειες)
<b>Συζήτηση</b>		
Περιορισμοί	20	Περιορισμοί της δοκιμής, αντιμετωπίζοντας τις πηγές πιθανών σφαλμάτων, την ανακρίβεια, και αν είναι σχετικό, την πολλαπλότητα των αναλύσεων
Γενικευσιμότητα	21	Γενικευσιμότητα (εξωτερική εγκυρότητα, εφαρμοσιμότητα) των ευρημάτων της δοκιμής
Ερμηνεία	22	Ερμηνεία συνεπής με τα αποτελέσματα, ισορροπώντας τα οφέλη και τις βλαβερές συνέπειες και λαμβάνοντας υπ' όψη άλλες σχετικές ενδείξεις
<b>Άλλες πληροφορίες</b>		
Καταχώρηση	23	Αριθμός καταχώρησης και όνομα του μητρώου δοκιμών
Πρωτόκολλο	24	Πού μπορεί να είναι προσβάσιμο το συνολικό πρωτόκολλο της δοκιμής, εάν είναι διαθέσιμο
Χρηματοδότηση	25	Πηγές χρηματοδότησης και άλλη υποστήριξη (όπως η προμήθεια φαρμάκων), ρόλος των χρηματοδοτών

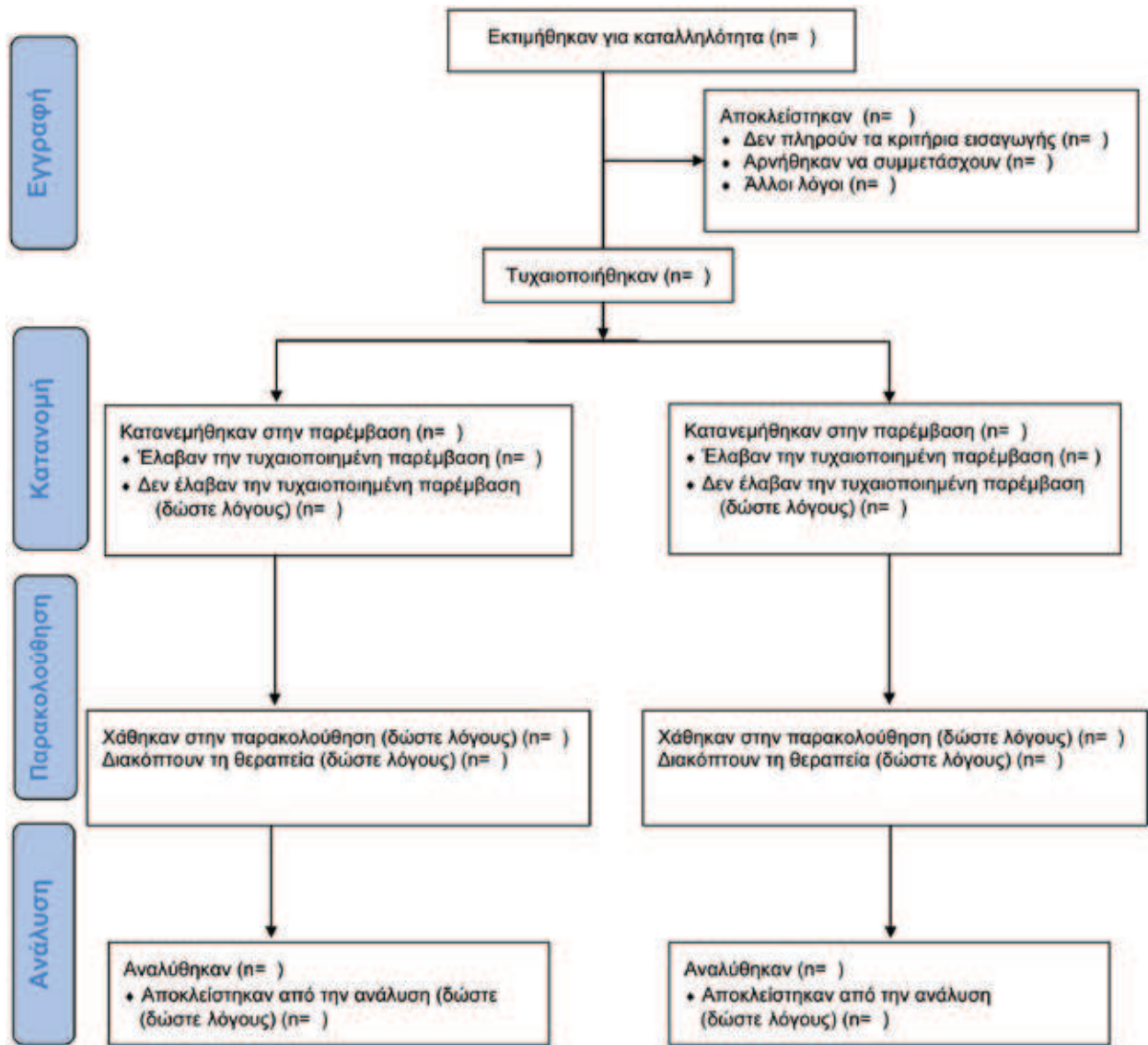
\*Προτείνουμε έντονα να διαβάζεται αυτή η δήλωση σε συνδυασμό με το CONSORT 2010, Εξήγηση και Επεξεργασία (CONSORT 2010 Explanation and Elaboration) για σημαντικές διευκρινίσεις σε όλα τα στοιχεία. Επίσης, αν είναι σχετικό, προτείνεται η ανάγνωση των επεκτάσεων του CONSORT για τυχαιοποιημένες δοκιμές συρροής,<sup>40</sup> δοκιμές μη υπεροχής και ισοδυναμίας,<sup>39</sup> μη φαρμακολογικές θεραπείες,<sup>43</sup> βοτανικές παρεμβάσεις<sup>44</sup> και πραγματικές δοκιμές (pragmatic trials).<sup>41</sup> Πρόσθετες επεκτάσεις θα είναι διαθέσιμες: Γι' αυτές και για άλλες επίκαιρες αναφορές σχετικές με αυτόν τον κατάλογο, ανατρέξτε στο δικτυακό τόπο [www.consort-statement.org](http://www.consort-statement.org)

#### 4. Η ΔΗΛΩΣΗ CONSORT 2010: ΕΠΕΞΗΓΗΣΗ ΚΑΙ ΕΠΕΞΕΡΓΑΣΙΑ

Κατά τη διάρκεια της αναθεώρησης του CONSORT του 2001, έγινε σαφές ότι η επεξήγηση και η επεξεργασία

των αρχών της δήλωσης CONSORT θα συνέδραμαν τους ερευνητές και άλλους να γράψουν ή να αξιολογήσουν τις δημοσιεύσεις/εργασίες των δοκιμών. Το άρθρο επεξήγησης και επεξεργασίας του CONSORT<sup>58</sup> δημοσιεύτηκε το 2001 παράλληλα με την έκδοση του 2001 της δήλωσης του





**Εικόνα 1.** Διάγραμμα ροής της προόδου των φάσεων μιας παράλληλης τυχαιοποιημένης δοκιμής δύο ομάδων (δηλαδή εγγραφή, διανομή παρέμβασης, παρακολούθηση και ανάλυση δεδομένων).<sup>52-54</sup>

CONSORT. Πραγματεύτηκε τη λογική και το επιστημονικό υπόβαθρο για κάθε στοιχείο και παρείχε δημοσιευμένα παραδείγματα καλής καταγραφής. Η λογική για την αναθεώρηση του συγκεκριμένου άρθρου είναι παρόμοια με αυτή της αναθεώρησης της δήλωσης, που περιγράφεται ανωτέρω. Παρακάτω περιγράφονται εν συντομία οι κύριες προσθήκες και οι διαγραφές στην εν λόγω έκδοση του άρθρου επεξήγησης και επεξεργασίας.

**5. ΕΠΕΞΗΓΗΣΗ ΚΑΙ ΑΝΑΠΤΥΞΗ ΤΟΥ CONSORT 2010: ΑΛΛΑΓΕΣ**

Έχουμε κάνει διάφορες ουσιαστικές, αλλά και μερικές αισθητικές αλλαγές σε αυτή την έκδοση του επεξηγημα-

τικού εγγράφου του CONSORT (πλήρεις λεπτομέρειες τονίζονται στην έκδοση του CONSORT του 2010<sup>59</sup>). Μερικές αντιπροσωπεύουν αλλαγές του καταλόγου του CONSORT. Υπάρχουν τρία νέα στοιχεία στον κατάλογο του CONSORT 2010, όπως το στοιχείο 24, το οποίο ζητά από τους συγγραφείς να καταγράψουν πού μπορεί να είναι προσβάσιμο το πρωτόκολλο της δοκιμής τους. Επίσης, έχουμε αναθεωρήσει μερικές υπάρχουσες επεξηγήσεις, περιλαμβανομένης της προσθήκης των πιο πρόσφατων αναφορών σε μεθοδολογικά δεδομένα, κι έχουμε χρησιμοποιήσει μερικά καλύτερα παραδείγματα. Έχουμε μεταφέρει το γλωσσάριο, το οποίο είναι τώρα διαθέσιμο στην ιστοσελίδα του CONSORT ([www.consort-statement.org](http://www.consort-statement.org)). Όπου είναι δυνατόν, περιγράφουμε τα ευρήματα των σχετικών εμπειρικών μελετών. Πολλά

έξοχα βιβλία για κλινικές δοκιμές προσφέρουν πληρέστερη συζήτηση των μεθοδολογικών ζητημάτων.<sup>60–62</sup> Τελικά, για ευκολία, μερικές φορές αναφερόμαστε σε «θεραπείες» και «ασθενείς», αν και αναγνωρίζουμε ότι όλες οι παρεμβάσεις που αξιολογούνται στις ΤΚΔ δεν είναι θεραπείες και όλοι οι συμμετέχοντες δεν είναι ασθενείς.

## 6. ΤΑ ΣΤΟΙΧΕΙΑ ΤΟΥ ΚΑΤΑΛΟΓΟΥ

### 6.1. Τίτλος και περίληψη

#### 6.1.1. Στοιχείο 1α. Αναγνώριση της δοκιμής ως τυχαιοποιημένη στον τίτλο

*Παράδειγμα:* «Μείωση καπνίσματος με εισπνευστικές συσκευές νικοτίνης: Διπλή-τυφλή, τυχαιοποιημένη κλινική δοκιμή αποτελεσματικότητας και ασφάλειας».<sup>63</sup>

*Εξήγηση:* Η δυνατότητα να αναγνωριστεί η δημοσίευση μιας τυχαιοποιημένης δοκιμής σε μια ηλεκτρονική βάση δεδομένων εξαρτάται σε μεγάλο βαθμό από τον τρόπο με τον οποίο καταχωρήθηκε. Οι καταχωρητές μπορεί να μην ταξινομήσουν μια δημοσίευση ως τυχαιοποιημένη δοκιμή εάν οι συγγραφείς δεν αναφέρουν σαφώς αυτή την πληροφορία.<sup>64</sup> Οι συγγραφείς, για να βοηθήσουν τη διασφάλιση ότι μια μελέτη είναι κατάλληλα καταχωρημένη και εύκολα αναγνωρίσιμη, πρέπει να χρησιμοποιούν τη λέξη «τυχαιοποιημένη» στον τίτλο για να δείξουν ότι οι συμμετέχοντες ταξινομήθηκαν τυχαία στις ομάδες σύγκρισής τους.

#### 6.1.2. Στοιχείο 1β. Δομημένη σύνοψη του σχεδιασμού, των μεθόδων, των αποτελεσμάτων και των συμπερασμάτων της δοκιμής. Αναφορικά με την ειδική καθοδήγηση, ανατρέξτε στο CONSORT για περιλήψεις<sup>45,65</sup>

*Εξήγηση:* Οι σαφείς, διαφανείς και επαρκώς λεπτομερείς περιλήψεις είναι σημαντικές, επειδή οι αναγνώστες βασίζονται συχνά στην αξιολόγηση μιας δοκιμής σε τέτοιες πληροφορίες. Μερικοί αναγνώστες χρησιμοποιούν μια περίληψη ως εργαλείο διαλογής για να αποφασίσουν εάν θα διαβάσουν το πλήρες άρθρο. Εν τούτοις, δεδομένου ότι όλες οι δοκιμές δεν είναι ελεύθερα διαθέσιμες και μερικοί επαγγελματίες υγείας δεν έχουν πρόσβαση στις πλήρεις δημοσιεύσεις των δοκιμών, οι αποφάσεις υγιονομικής περιθάλψης λαμβάνονται μερικές φορές με βάση τις περιλήψεις των τυχαιοποιημένων κλινικών δοκιμών.<sup>66</sup>

Μια περίληψη περιοδικού πρέπει να περιέχει επαρκείς πληροφορίες για μια δοκιμή προκειμένου να χρησιμεύσει ως ένα ακριβές αρχείο της διεξαγωγής και των ευρημάτων

της, παρέχοντας τις βέλτιστες πληροφορίες για τη δοκιμή μέσα στους περιορισμούς του διαστήματος και της διάταξης ενός περιοδικού. Μια κατάλληλα δομημένη και γραμμένη περίληψη βοηθά τα άτομα να αξιολογήσουν γρήγορα τη σχετικότητα των ευρημάτων, ενώ παράλληλα συνδράμει στην ανάκτηση των σχετικών δημοσιεύσεων από ηλεκτρονικές βάσεις δεδομένων.<sup>67</sup> Η περίληψη πρέπει να απεικονίζει επακριβώς τι περιλαμβάνεται στο πλήρες άρθρο του περιοδικού και δεν πρέπει να περιέχει πληροφορίες που δεν εμφανίζονται στο σώμα του εγγράφου. Μελέτες που συνέκριναν την ακρίβεια των πληροφοριών που αναφέρονται σε μια περίληψη περιοδικού με αυτή που παρατίθεται στο κείμενο της πλήρους δημοσίευσης έχουν βρει ισχυρισμούς που είναι αντιφατικοί, ή ελλείπουν, από το σώμα του πλήρους άρθρου.<sup>68–71</sup> Αντίθετα, η παράλειψη σημαντικών βλαβερών συνεπειών από την περίληψη θα μπορούσε να παραπλανήσει σοβαρά την ερμηνεία κάποιου για τα ευρήματα της δοκιμής.<sup>42,72</sup>

Μια πρόσφατη επέκταση στη δήλωση του CONSORT παρέχει έναν κατάλογο βασικών στοιχείων που πρέπει να περιλάβουν οι συγγραφείς κατά την αναφορά των κύριων αποτελεσμάτων μιας τυχαιοποιημένης δοκιμής σε μια περίληψη περιοδικού (ή συνεδρίου) (πίν. 2).<sup>45</sup> Συστήνεται έντονα η χρήση δομημένων περιλήψεων για την καταγραφή των τυχαιοποιημένων δοκιμών. Οι περιλήψεις αυτές παρέχουν στους αναγνώστες τις πληροφορίες για τη δοκιμή κάτω από μια σειρά επικεφαλίδων σχετικά με το σχεδιασμό, τη διεξαγωγή, την ανάλυση και την ερμηνεία.<sup>73</sup> Σε μερικές μελέτες έχει βρεθεί ότι οι δομημένες περιλήψεις είναι υψηλότερης ποιότητας από τις περισσότερο παραδοσιακές περιγραφικές περιλήψεις<sup>74,75</sup> και ότι επιτρέπουν στους αναγνώστες να βρουν τις πληροφορίες πιο εύκολα.<sup>76</sup> Αναγνωρίζουμε ότι πολλά περιοδικά έχουν αναπτύξει τη δική τους δομή και το όριο λέξεων για την αναφορά των περιλήψεων. Δεν είναι πρόθεσή μας να προτείνουμε μεταβολές σε αυτά τα σχήματα, αλλά να συστήσουμε ποιες πληροφορίες πρέπει να αναφερθούν.

### 6.2. Εισαγωγή

#### 6.2.1. Στοιχείο 2α. Επιστημονικό υπόβαθρο και εξήγηση της λογικής

*Παράδειγμα:* «Η χειρουργική επέμβαση είναι θεραπεία εκλογής για τους ασθενείς με μη μικροκυτταρικό καρκίνο του πνεύμονα (ΜΜΚΠ), σταδίου Ι και ΙΙ... Μια μετα-ανάλυση του ΜΜΚΠ συνδύασε τα αποτελέσματα από 8 τυχαιοποιημένες δοκιμές της χειρουργικής επέμβασης εναντίον της χειρουργικής επέμβασης και της βοηθητικής χημειοθεραπείας βασισμένη στη σισπλατίνη και παρουσίασε ένα μικρό, αλλά όχι σημαντικό ( $p=0,08$ ), απόλυτο όφελος επιβίωσης

**Πίνακας 2.** Στοιχεία που θα πρέπει να περιλαμβάνονται κατά τη δημοσίευση μιας τυχαιοποιημένης δοκιμής σε μια περίληψη περιοδικού.

Στοιχείο	Περιγραφή
Συγγραφείς	Λεπτομέρειες επικοινωνίας για το συγγραφέα αλληλογραφίας
Σχεδιασμός δοκιμής	Περιγραφή του σχεδιασμού της δοκιμής (όπως παράλληλος, συρροής, μη υπεροχής)
Μέθοδοι	
Συμμετέχοντες	Κριτήρια καταλληλότητας των συμμετεχόντων και οι συνθήκες όπου συλλέχθηκαν τα δεδομένα
Παρεμβάσεις	Οι παρεμβάσεις που προορίζονται για κάθε ομάδα
Στόχοι	Συγκεκριμένοι στόχοι ή υποθέσεις
Έκβαση	Σαφώς καθορισμένη κύρια έκβαση γι' αυτή την εργασία
Τυχαιοποίηση	Πώς οι συμμετέχοντες διανεμήθηκαν στις παρεμβάσεις
Τυφλοποίηση (αποκρυπτοποίηση)	Εάν οι συμμετέχοντες, οι προμηθευτές φροντίδας και εκείνοι που αξιολογούν τις εκβάσεις τυφλοποιήθηκαν στην ταξινόμηση σε ομάδες
Αποτελέσματα	
Αριθμοί που τυχαιοποιήθηκαν	Ο αριθμός των συμμετεχόντων που τυχαιοποιήθηκαν σε κάθε ομάδα
Στρατολόγηση	Κατάσταση της δοκιμής
Αριθμοί που αναλύθηκαν	Ο αριθμός των συμμετεχόντων που αναλύθηκε σε κάθε ομάδα
Έκβαση	Για την κύρια έκβαση, ένα αποτέλεσμα για κάθε ομάδα και το εκτιμώμενο μέγεθος επίδρασης και η ακρίβειά του
Βλαβερές συνέπειες	Σημαντικές ανεπιθύμητες ενέργειες ή παρενέργειες
Συμπεράσματα	Γενική ερμηνεία των αποτελεσμάτων
Καταχώρηση δοκιμής	Αριθμός καταχώρησης και όνομα του μητρώου εγγραφής της δοκιμής
Χρηματοδότηση	Πηγή χρηματοδότησης

περίπου 5% σε 5 έτη (από 50% σε 55%). Την περίοδο κατά την οποία σχεδιάστηκε η τρέχουσα δοκιμή (μέσα του 1990) η επικουρική χημειοθεραπεία δεν είχε καταστεί η καθιερωμένη κλινική πρακτική... Η κλινική αιτιολογία της νέας βοηθητικής χημειοθεραπείας είναι τριπλή: Η υποτροπή του πρωτογενούς καρκίνου θα μπορούσε να επιτευχθεί έτσι, διευκολύνοντας και απλουστεύοντας ή μειώνοντας την επόμενη χειρουργική επέμβαση. Οι μη ανιχνεύσιμες μικρομεταστάσεις θα μπορούσαν να έχουν διαχειριστεί στην έναρξη της θεραπείας και ενδέχεται να υπάρξει αναστολή του πιθανού ερεθίσματος στον υπολειπόμενο καρκίνο μέσω αυξητικών παραγόντων που εκλύονται από τη χειρουργική επέμβαση και από την επακόλουθη επούλωση του τραύματος... Επομένως, η τρέχουσα δοκιμή ήταν οργανωμένη για να συγκρίνει, σε ασθενείς με εξαιρεσιμο ΜΜΚΠ, τη χειρουργική επέμβαση μόνη εναντίον τριών κύκλων της βασικής χημειοθεραπείας με πλατίνη που ακολουθείται από χειρουργική παρέμβαση, όσον αφορά στη γενική επιβίωση, στην ποιότητα ζωής, στη σταδιοποίηση, στα ποσοστά εξαιρεσιμότητας, στην επέκταση της χειρουργικής επέμβασης, καθώς και στο χρόνο και στην περιοχή της υποτροπής».<sup>77</sup>

*Εξήγηση:* Χαρακτηριστικά, η εισαγωγή αποτελείται από κείμενο ελεύθερης ροής, στο οποίο οι συγγραφείς εξηγούν το επιστημονικό υπόβαθρο και τη λογική της δοκιμής τους, καθώς και τη γενική περιγραφή της. Επίσης, μπορεί να

είναι πρόπον να περιληφθούν εδώ οι στόχοι της δοκιμής (βλέπε στοιχείο 2β). Η λογική μπορεί να είναι επεξηγηματική (για παράδειγμα, προκειμένου να αξιολογήσουν την πιθανή επίδραση ενός φαρμάκου στη νεφρική λειτουργία) ή πραγματική (παραδείγματος χάρη, για να καθοδηγήσει την πρακτική με τη σύγκριση των οφελών και των βλαβερών συνεπειών δύο θεραπειών). Οι συγγραφείς πρέπει να αναφέρουν οποιοσδήποτε ενδείξεις των οφελών και των βλαβερών συνεπειών των ενεργών παρεμβάσεων που περιλαμβάνονται σε μια δοκιμή και πρέπει να προτείνουν μια εύλογη εξήγηση για το πώς μπορούν να λειτουργήσουν οι παρεμβάσεις, εάν αυτό δεν είναι προφανές.<sup>78</sup>

Η Διακήρυξη του Helsinki δηλώνει ότι η βιοϊατρική έρευνα που περιλαμβάνει ανθρώπους πρέπει να βασίζεται στη λεπτομερή γνώση της επιστημονικής βιβλιογραφίας.<sup>79</sup> Δηλαδή, είναι ανήθικο να εκτίθενται οι άνθρωποι άσκοπα στους κινδύνους της έρευνας. Μερικές κλινικές δοκιμές έχουν αποδειχθεί ότι είναι άσκοπες, επειδή η υπόθεση που εξέτασαν ήταν ήδη ή μπορούσε να έχει απαντηθεί από μια συστηματική ανασκόπηση της υπάρχουσας βιβλιογραφίας.<sup>80,81</sup> Κατά συνέπεια, η ανάγκη για μια νέα δοκιμή πρέπει να δικαιολογηθεί στην εισαγωγή. Ιδανικά, θα πρέπει να περιλαμβάνει μια αναφορά σε μια συστηματική ανασκόπηση προηγούμενων παρόμοιων δοκιμών ή μια σημείωση της απουσίας τέτοιων δοκιμών.<sup>82</sup>

### 6.2.2. Στοιχείο 2β. Ειδικοί στόχοι ή υποθέσεις

*Παράδειγμα:* «Στην παρούσα μελέτη εξετάσαμε την υπόθεση ότι η πολιτική ενεργούς διαχείρισης των πόνων του τοκετού θα μπορούσε (α) να ελαττώσει το ποσοστό της καισαρικής τομής, (β) να μειώσει το ποσοστό των παρατεταμένων τοκετών και (γ) να μην επηρεάσει τη μητρική ικανοποίηση από την εμπειρία της γέννας». <sup>83</sup>

*Εξήγηση:* Οι στόχοι είναι τα ερωτήματα για τα οποία η δοκιμή σχεδιάστηκε προκειμένου να δώσει απαντήσεις. Συχνά, σχετίζονται με την αποτελεσματικότητα μιας συγκεκριμένης θεραπευτικής ή προληπτικής παρέμβασης. Οι υποθέσεις είναι προκαθορισμένα ερωτήματα που εξετάζονται ώστε να συνδράμουν στην προσέγγιση των στόχων. Οι υποθέσεις είναι πιο συγκεκριμένες από τους στόχους και υπόκεινται σε λεπτομερή στατιστική αξιολόγηση. Στην πράξη, οι στόχοι και οι υποθέσεις δεν είναι πάντα εύκολα διαχωρίσιμοι. Οι περισσότερες δημοσιεύσεις των ΤΚΔ παρέχουν επαρκείς πληροφορίες για τους στόχους και τις υποθέσεις της δοκιμής. <sup>84</sup>

## 6.3. Μέθοδοι

### 6.3.1. Στοιχείο 3α. Περιγραφή του σχεδιασμού της δοκιμής (όπως παράλληλος, παραγοντικός, περιλαμβανομένης της αναλογίας διανομής)

*Παράδειγμα:* «Αυτή ήταν μια πολυκεντρική, διαστρωματωμένη (ηλικίες 6–11 ετών και 12–17 ετών, με ανισομερή τυχαίοποίηση [2:1]), διπλή-τυφλή, ελεγχόμενη με εικονικό φάρμακο, παράλληλων ομάδων μελέτη που πραγματοποιήθηκε στις Ηνωμένες Πολιτείες (41 περιοχές)». <sup>85</sup>

*Εξήγηση:* Η λέξη «σχεδιασμός» χρησιμοποιείται συχνά για να αναφερθεί σε όλες τις πλευρές του πώς μια δοκιμή είναι οργανωμένη, αλλά έχει επίσης μια στενότερη ερμηνεία. Πολλές συγκεκριμένες πτυχές του ευρύτερου σχεδιασμού της δοκιμής, περιλαμβανομένων των λεπτομερειών της τυχαίοποίησης και της τυφλοποίησης, εξετάζονται αλλού στον κατάλογο του CONSORT. Εδώ, αναζητούμε πληροφορίες για τον τύπο της δοκιμής, όπως παράλληλων ομάδων ή παραγοντική, και το εννοιολογικό πλαίσιο, όπως ανωτερότητας ή μη υπεροχής, καθώς και άλλα σχετικά ζητήματα που δεν αντιμετωπίζονται αλλού στον κατάλογο.

Η δήλωση CONSORT εστιάζει κυρίως σε δοκιμές με συμμετέχοντες τυχαίοποιημένους ξεχωριστά στη μία από τις δύο «παράλληλες» ομάδες. Στην πραγματικότητα, λίγο περισσότερες από τις μισές δημοσιευμένες δοκιμές έχουν έναν τέτοιο σχεδιασμό. <sup>16</sup> Οι κύριοι εναλλακτικοί σχεδιασμοί είναι παράλληλοι πολλαπλών σκελών (multi-arm), διασταυ-

ρούμενοι, συρροής <sup>40</sup> και παραγοντικοί σχεδιασμοί. Επίσης, οι περισσότερες δοκιμές στοχεύουν να αναγνωρίσουν την ανωτερότητα μιας νέας παρέμβασης, εάν υπάρχει, αλλά άλλες σχεδιάστηκαν για να αξιολογήσουν τη μη υπεροχή ή την ισοδυναμία. <sup>39</sup> Είναι σημαντικό ότι οι ερευνητές περιγράφουν σαφώς αυτές τις πτυχές της δοκιμής τους, περιλαμβανομένης της μονάδας της τυχαίοποίησης (όπως ασθενής, ιατρείο Γενικής Ιατρικής, κάκωση). Είναι επιθυμητό επίσης να περιληφθούν αυτές οι λεπτομέρειες στην περίληψη (βλέπε στοιχείο 1β).

Εάν υιοθετείται ένας λιγότερο συνήθης σχεδιασμός, οι συγγραφείς ενθαρρύνονται να εξηγήσουν την επιλογή τους, ειδικά, δεδομένου ότι τέτοιοι σχεδιασμοί μπορούν να υποδηλώσουν την ανάγκη για ένα μεγαλύτερο μέγεθος δείγματος ή μια πιο σύνθετη ανάλυση και ερμηνεία.

Αν και οι περισσότερες δοκιμές χρησιμοποιούν ίση τυχαίοποίηση (όπως η 1:1 για δύο ομάδες), είναι χρήσιμο να παρέχει σαφώς την αναλογία διανομής. Για τις δοκιμές φαρμάκων, ο προσδιορισμός της φάσης της δοκιμής (I–IV) μπορεί επίσης να είναι σχετικός.

### 6.3.2. Στοιχείο 3β. Σημαντικές αλλαγές στις μεθόδους ύστερα από την έναρξη της δοκιμής (όπως στα κριτήρια καταλληλότητας), με αιτιολόγηση

*Παράδειγμα:* «Οι ασθενείς ταξινομήθηκαν τυχαία σε μία από τις έξι παράλληλες ομάδες, αρχικά σε αναλογία 1:1:1:1:1:1, για να λάβουν είτε το ένα από τα πέντε otaμixaban... θεραπευτικά σχήματα... είτε έναν ενεργό έλεγχο της ηπαρίνης... Μια ανεξάρτητη Επιτροπή Επιτήρησης Δεδομένων (Data Monitoring Committee) επανεξέτασε τα μη τυφλοποιημένα δεδομένα για την ασφάλεια του ασθενούς, ενώ δεν πραγματοποιήθηκε κάποια ενδιάμεση ανάλυση για την αποτελεσματικότητα ή τη ματαιότητα. Κατά τη διάρκεια της δοκιμής, αυτή η Επιτροπή συνέστησε ότι η ομάδα που ελάμβανε τη χαμηλότερη δόση otaμixaban (0,035 mg/kg/ώρα) διακόπτεται λόγω κλινικών ενδείξεων για ανεπαρκή αντιπηκτικότητα. Το πρωτόκολλο τροποποιήθηκε αμέσως, σύμφωνα με εκείνη τη σύσταση, και οι συμμετέχοντες στη συνέχεια ταξινομήθηκαν τυχαία σε αναλογία 2:2:2:2:1 στην παραμονή στο otaμixaban και στις ομάδες ελέγχου, αντίστοιχα». <sup>86</sup>

*Εξήγηση:* Λίγες δοκιμές μπορούν να αρχίσουν χωρίς κάποιο αμετάβλητο σχεδιασμό (δηλαδή, να είναι εξ ολοκλήρου διερευνητικός), ενώ οι περισσότερες θα έχουν ένα πρωτόκολλο που διευκρινίζει με μεγάλη λεπτομέρεια πώς θα πραγματοποιηθεί η δοκιμή. Μπορεί να υπάρξουν αποκλίσεις από το αρχικό πρωτόκολλο, καθώς είναι αδύνατο να προβλεφθεί κάθε πιθανή μεταβολή στις συνθήκες κατά τη διάρκεια μιας δοκιμής. Μερικές δοκιμές επομένως



έχουν σημαντικές αλλαγές στις μεθόδους, μετά από την έναρξη της δοκιμής.

Οι τροποποιήσεις θα μπορούσαν να οφείλονται στις εξωτερικές πληροφορίες που διατίθενται από άλλες μελέτες, ή σε εσωτερικές οικονομικές δυσχέρειες, ή θα μπορούσαν να οφείλονται σε ένα απογοητευτικό ποσοστό στρατολόγησης. Τέτοιες αλλαγές πρωτοκόλλου πρέπει να εφαρμόζονται χωρίς άρση της τυφλοποίησης στα συσσωρευμένα δεδομένα των εκβάσεων των συμμετεχόντων. Σε μερικές δοκιμές, μια ανεξάρτητη επιτροπή επιτήρησης των δεδομένων θα έχει ως τμήμα της αποστολής της την πιθανότητα σύστασης αλλαγών στο πρωτόκολλο βλέποντας τα μη τυφλοποιημένα δεδομένα. Τέτοιες μεταβολές μπορούν να επηρεάσουν τις μεθόδους της μελέτης (όπως αλλαγές στα θεραπευτικά σχήματα, στα κριτήρια καταλληλότητας, στην αναλογία τυχαιοποίησης ή στη διάρκεια της παρακολούθησης) ή τη διενέργεια της δοκιμής (όπως η απόρριψη ενός κέντρου με κακή ποιότητα δεδομένων).<sup>87</sup>

Μερικές δοκιμές είναι οργανωμένες με έναν επίσημο «προσαρμοστικό» σχεδιασμό. Παγκόσμια, δεν υπάρχει κανένας αποδεκτός ορισμός αυτών των σχεδιασμών, αλλά ένας λειτουργικός ορισμός μπορεί να είναι «ένας σχεδιασμός μελέτης πολλών σταδίων που χρησιμοποιεί συσσωρευμένα δεδομένα για να αποφασίσει πώς να τροποποιήσει τις πτυχές της μελέτης, χωρίς την υπονόμηση της εγκυρότητας και της ακεραιότητας της δοκιμής».<sup>88</sup> Οι τροποποιήσεις είναι συνήθως στα μεγέθη των δειγμάτων και στον αριθμό των θεραπευτικών σκελών και μπορεί να οδηγήσουν σε αποφάσεις που λαμβάνονται γρηγορότερα και με περισσότερο αποδοτική χρήση των πόρων. Υπάρχουν, εν τούτοις, σημαντικά ηθικά, στατιστικά και πρακτικά ζητήματα τα οποία πρέπει να ληφθούν υπ' όψη σ' έναν τέτοιο σχεδιασμό.<sup>89,90</sup>

Εάν οι τροποποιήσεις είναι σαφώς τμήμα του σχεδιασμού της δοκιμής ή σε απάντηση των μεταβαλλόμενων συνθηκών, είναι ουσιαστικό να αναφέρονται πλήρως για να βοηθήσουν τον αναγνώστη να ερμηνεύσει τα αποτελέσματα. Οι μεταβολές των πρωτοκόλλων αυτή την περίοδο δεν είναι καλά αναφερόμενες. Μια ανασκόπηση των συγκρίσεων με πρωτόκολλα έδειξε ότι περίπου τα μισά από τα άρθρα περιοδικών που περιγράφουν ΤΚΔ είχαν μια ανεξήγητη απόκλιση στις κύριες εκβάσεις.<sup>57</sup> Συχνά, ανεξήγητες αποκλίσεις έχουν επίσης παρατηρηθεί για τις λεπτομέρειες της τυχαιοποίησης, της τυφλοποίησης και των στατιστικών αναλύσεων.<sup>92</sup>

### 6.3.3. Στοιχείο 4α. Κριτήρια καταλληλότητας για τους συμμετέχοντες

*Παράδειγμα:* «Κατάλληλοι συμμετέχοντες ήταν όλοι οι

ενήλικες, ηλικίας 18 ετών ή περισσότερο, με HIV-λοίμωξη που ικανοποιούσαν τα κριτήρια καταλληλότητας για την αντιρετροϊκή θεραπεία σύμφωνα με τις εθνικές οδηγίες θεραπείας HIV του Malawi (κλινικό στάδιο III ή IV κατά WHO ή οποιοδήποτε στάδιο με αριθμό CD4 <250/mm<sup>3</sup>) και οι οποίοι άρχιζαν θεραπεία με έναν BMI <18,5. Τα κριτήρια αποκλεισμού ήταν η εγκυμοσύνη και ο θηλασμός ή η συμμετοχή σε ένα άλλο συμπληρωματικό πρόγραμμα σίτισης».<sup>93</sup>

*Εξήγηση:* Απαιτείται μια περιεκτική περιγραφή των κριτηρίων καταλληλότητας που χρησιμοποιούνται για να επιλέξουν τους συμμετέχοντες της δοκιμής προκειμένου να βοηθήσει τους αναγνώστες να ερμηνεύσουν τη μελέτη. Ιδιαίτερα, η σαφής κατανόηση αυτών των κριτηρίων είναι ένα από τα διάφορα δεδομένα που απαιτούνται για να κρίνουν σε ποιον εφαρμόζονται τα αποτελέσματα μιας δοκιμής, που είναι η δυνατότητα γενίκευσης της δοκιμής (εφαρμοσιμότητα) και η σχετικότητα με την κλινική ή την πρακτική δημόσιας υγείας (βλέπε στοιχείο 21).<sup>94</sup> Μια περιγραφή της μεθόδου στρατολόγησης, όπως από παραπομπή ή αυτο-επιλογή (για παράδειγμα, μέσω των διαφημίσεων), αποτελεί επίσης σημαντικό στοιχείο σε αυτό το πλαίσιο. Επειδή εφαρμόζονται πριν από την τυχαιοποίηση, τα κριτήρια καταλληλότητας δεν επηρεάζουν την εσωτερική εγκυρότητα μιας δοκιμής, αλλά είναι κεντρικά στην εξωτερική εγκυρότητά της.

Τα χαρακτηριστικά και ευρέως αποδεκτά κριτήρια καταλληλότητας αφορούν στη φύση και στο στάδιο της νόσου που μελετάται, στον αποκλεισμό των προσώπων που είναι ιδιαίτερα ευάλωτα στις βλαβερές συνέπειες από την παρέμβαση της μελέτης και στα ζητήματα που απαιτούνται για να εξασφαλίσουν ότι η μελέτη ικανοποιεί τους νομικούς και τους ηθικούς κανόνες. Η μετά από ενημέρωση συγκατάθεση των συμμετεχόντων, για παράδειγμα, συνήθως απαιτείται στις μελέτες παρέμβασης. Η συχνή διάκριση μεταξύ των κριτηρίων εισόδου και αποκλεισμού είναι περιττή. Το ίδιο κριτήριο μπορεί να διατυπωθεί για να περιλάβει ή να αποκλείσει συμμετέχοντες.<sup>95</sup>

Παρά τη σημαντικότητά τους, τα κριτήρια καταλληλότητας συχνά δεν αναφέρονται επαρκώς. Για παράδειγμα, 8 δημοσιευμένες δοκιμές που οδηγούν σε κλινικές ειδοποιήσεις από τα εθνικά ιδρύματα υγείας καθόρισαν ένα μέσο όρο 31 κριτηρίων καταλληλότητας στα πρωτόκολλά τους, αλλά μόνο 63% των κριτηρίων αναφέρθηκαν σε άρθρα περιοδικών και μόνο 19% αναφέρθηκαν σε κλινικές ειδοποιήσεις.<sup>96</sup> Παρόμοιες ανεπάρκειες βρέθηκαν στις κλινικές δοκιμές του HIV.<sup>97</sup> Μεταξύ 364 δημοσιεύσεων των ΤΚΔ στη Χειρουργική, το 25% δεν διευκρίνισε οποιαδήποτε κριτήρια καταλληλότητας.<sup>98</sup>

#### 6.3.4. Στοιχείο 4β. Συνθήκες και τόποι στους οποίους συλλέχθηκαν τα δεδομένα

*Παράδειγμα:* «Η μελέτη πραγματοποιήθηκε στην κλινική αντι-ιικής/αντιρετροϊκής θεραπείας του νοσοκομείου Queen Elizabeth Central στην Blantyre του Malawi, από τον Ιανουάριο του 2006 έως τον Απρίλιο του 2007. Η Blantyre είναι η μεγαλύτερη εμπορική πόλη του Malawi, με πληθυσμό 1.000.000 κατοίκους και έναν εκτιμώμενο επιπολασμό HIV-λοίμωξη 27% στους ενήλικες το 2004». <sup>93</sup>

*Εξήγηση:* Μαζί με τα κριτήρια καταλληλότητας για τους συμμετέχοντες (βλέπε στοιχείο 4α) και την περιγραφή των παρεμβάσεων (βλέπε στοιχείο 5), οι πληροφορίες για τις συνθήκες και τους τόπους είναι κρίσιμες για να κρίνουν την εφαρμοσιμότητα και τη γενικευσιμότητα μιας δοκιμής. Οι συμμετέχοντες στρατολογήθηκαν από την πρωτοβάθμια, τη δευτεροβάθμια ή την τριτοβάθμια φροντίδα υγείας ή από την κοινότητα; Τα υγειονομικά ιδρύματα ποικίλλουν πολύ στην οργάνωσή τους, την εμπειρία, καθώς και τους πόρους και το βασικό κίνδυνο για τις συνθήκες κάτω από τις οποίες γίνεται η έρευνα. Άλλες πτυχές των συνθηκών (περιλαμβανομένου του κοινωνικού, του οικονομικού και του πολιτιστικού περιβάλλοντος και του κλίματος) μπορεί επίσης να επηρεάσουν την εξωτερική εγκυρότητα μιας μελέτης.

Οι συγγραφείς πρέπει να αναφέρουν τον αριθμό και τον τύπο των συνθηκών, καθώς και να περιγράψουν τους εμπλεκόμενους προμηθευτές φροντίδας υγείας. Πρέπει να αναφέρουν τους τόπους στους οποίους διεξήχθη η μελέτη, περιλαμβανομένης της χώρας, της πόλης εάν είναι εφαρμόσιμο, και του άμεσου περιβάλλοντος (για παράδειγμα, την κοινότητα, το ιατρείο, την κλινική των νοσοκομείων ή τη μονάδα ασθενών). Ειδικότερα, πρέπει να είναι σαφές εάν η δοκιμή πραγματοποιήθηκε σε ένα ή περισσότερα κέντρα («πολυκεντρικές δοκιμές»). Αυτή η περιγραφή πρέπει να παρέχει αρκετές πληροφορίες, έτσι ώστε οι αναγνώστες να μπορούν να κρίνουν εάν τα αποτελέσματα της δοκιμής θα μπορούσαν να είναι σχετικά με τις δικές τους συνθήκες. Το περιβάλλον στο οποίο πραγματοποιείται η δοκιμή μπορεί να διαφέρει σημαντικά από τις συνθήκες στις οποίες τα αποτελέσματα της δοκιμής χρησιμοποιούνται αργότερα για να καθοδηγήσουν την πρακτική και την πολιτική. <sup>94,99</sup> Οι συγγραφείς πρέπει επίσης να αναφέρουν οποιεσδήποτε άλλες πληροφορίες για τις συνθήκες και τους τόπους που θα μπορούσαν να έχουν επηρεάσει τα παρατηρούμενα αποτελέσματα, όπως προβλήματα στη μετακίνηση που μπορεί να επηρεάζουν τη συμμετοχή των ασθενών ή καθυστερήσεις στη διαχείριση των παρεμβάσεων.

#### 6.3.5. Στοιχείο 5. Οι παρεμβάσεις της κάθε ομάδας με αρκετή λεπτομέρεια έτσι ώστε να είναι εφικτή η επανάληψη, περιλαμβανομένου του πώς και πότε πραγματικά χορηγήθηκαν

*Παραδείγματα:* «Στην POISE, οι ασθενείς έλαβαν την πρώτη δόση του φαρμάκου μελέτης (δηλαδή από του στόματος μετοπρολόλη 100 mg παρατεταμένης αποδέσμευσης ή αντίστοιχο εικονικό φάρμακο) 2–4 ώρες πριν από τη χειρουργική επέμβαση. Η χορήγηση του φαρμάκου της μελέτης απαιτούσε έναν καρδιακό ρυθμό περίπου 50 bpm ή περισσότερο και μια συστολική αρτηριακή πίεση 100 mmHg ή μεγαλύτερη. Οι εν λόγω αιμοδυναμικές παράμετροι ελέγχονταν πριν από κάθε χορήγηση.

Εάν, οποιαδήποτε στιγμή κατά τη διάρκεια των πρώτων 6 ωρών μετά από τη χειρουργική επέμβαση, ο καρδιακός ρυθμός ήταν 80 bpm ή μεγαλύτερος και η συστολική αρτηριακή πίεση 100 mmHg ή υψηλότερη, οι ασθενείς ελάμβαναν την πρώτη μετεγχειρητική δόση τους (μετοπρολόλη παρατεταμένης αποδέσμευσης 100 mg ή αντίστοιχο εικονικό φάρμακο) από το στόμα. Εάν το φάρμακο μελέτης δεν χορηγείτο κατά τη διάρκεια των πρώτων 6 ωρών, οι ασθενείς ελάμβαναν την πρώτη μετεγχειρητική δόση τους 6 ώρες μετά από τη χειρουργική επέμβαση. Δώδεκα ώρες μετά από την πρώτη μετεγχειρητική δόση, οι ασθενείς άρχισαν από το στόμα μετοπρολόλη παρατεταμένης αποδέσμευσης 200 mg ή εικονικό φάρμακο καθημερινά για 30 ημέρες. Εάν ο καρδιακός ρυθμός ενός ασθενούς ήταν συνέχεια <45 bpm ή η συστολική αρτηριακή πίεσή του <100 mmHg, το φάρμακο μελέτης διακοπτόταν έως ότου ανακτηθούν ο καρδιακός ρυθμός ή η συστολική αρτηριακή πίεση. Το φάρμακο μελέτης επαναχορηγείτο στη συνέχεια σε δόση 100 mg μία φορά ημερησίως. Οι ασθενείς των οποίων ο καρδιακός ρυθμός ήταν σταθερά 45–49 bpm και η συστολική αρτηριακή πίεση υπερέβαινε τα 100 mm Hg έπαιρναν το φάρμακο της μελέτης καθυστερημένα για 12 ώρες». <sup>100</sup>

«Οι ασθενείς ταξινομήθηκαν τυχαία για να λάβουν έναν επί παραγγελία νάρθηκα νεοπρενίου που φοριέται τη νύκτα ή τη συνηθισμένη φροντίδα. Ο νάρθηκας ήταν λειτουργικά άκαμπος, με ένδειξη για χρήση μόνο τη νύκτα. Κάλυπτε τη βάση του αντίχειρα και του θέναρος αλλά όχι της πηχεοκαρπικής άρθρωσης. Οι νάρθηκες κατασκευάστηκαν από 3 εκπαιδευμένους εργοθεραπευτές, οι οποίοι ρύθμισαν το νάρθηκα για κάθε ασθενή έτσι ώστε το πρώτο πλέγμα να μπορεί να ανοιχθεί και ο αντίχειρας να τοποθετηθεί σε αντίθεση με το πρώτο μακρύ δάκτυλο. Οι ασθενείς ενθαρρύνθηκαν για να έλθουν σε επαφή με τον εργοθεραπευτή εάν θεωρούσαν ότι ο νάρθηκας χρειαζόταν ρύθμιση, αν ο πόνος αυξανόταν φορώντας το νάρθηκα ή

αν είχαν ανεπιθύμητες επιδράσεις (όπως η διάβρωση του δέρματος). Επειδή καμιά θεραπεία δεν μπορεί να θεωρηθεί ο χρυσός κανόνας σε αυτή την κατάσταση, οι ασθενείς και στην ομάδα ελέγχου και στην ομάδα παρέμβασης έλαβαν τη συνηθισμένη φροντίδα κατά την κρίση του ιατρού τους (γενικού ιατρού ή ρευματολόγου). Αποφασίσαμε να μη χρησιμοποιήσουμε μια εικονική παρέμβαση επειδή, απ' ό,τι είναι γνωστό, καμιά εικονική παρέμβαση για την τοποθέτηση νάρθηκα δεν έχει κατορθώσει την επιτυχή τυφλοποίηση των ασθενών, όπως συστήνεται». <sup>101</sup>

*Εξήγηση:* Οι συγγραφείς πρέπει να περιγράψουν κάθε παρέμβαση λεπτομερώς, περιλαμβανομένων των παρεμβάσεων ελέγχου. Η περιγραφή πρέπει να επιτρέπει σε έναν κλινικό ιατρό που θέλει να χρησιμοποιήσει την παρέμβαση να γνωρίζει ακριβώς πώς να διαχειριστεί την παρέμβαση που αξιολογήθηκε στη δοκιμή. <sup>102</sup> Για μια φαρμακευτική παρέμβαση, οι πληροφορίες θα πρέπει να περιλαμβάνουν το όνομα του φαρμάκου, τη δόση, την οδό χορήγησης (όπως στοματική, ενδοφλέβια), το χρόνο και τη διάρκεια της χορήγησης, τις συνθήκες κάτω από τις οποίες έγιναν οι παρεμβάσεις και την τιτλοποίηση της θεραπευτικής αγωγής εφ' όσον είναι εφαρμόσιμη. Εάν η ομάδα ελέγχου πρόκειται να λάβει τη «συνηθισμένη φροντίδα», είναι σημαντικό να περιγράψει λεπτομερώς τι περιλαμβάνει εκείνη. Εάν η ομάδα ελέγχου ή η ομάδα παρέμβασης πρόκειται να λάβει ένα συνδυασμό παρεμβάσεων, οι συγγραφείς πρέπει να παρέχουν μια λεπτομερή περιγραφή κάθε παρέμβασης, μια επεξήγηση της σειράς στην οποία εισάγεται ή αποσύρεται ο συνδυασμός των παρεμβάσεων, καθώς και οι αφορμές για την εισαγωγή τους εφ' όσον είναι εφαρμόσιμες.

Συγκεκριμένες επεκτάσεις της δήλωσης CONSORT εξετάζουν την αναφορά των μη φαρμακολογικών και των παρεμβάσεων με βότανα, καθώς και των ιδιαίτερων απαιτήσεων καταγραφής τους (όπως οι ειδικές γνώσεις, λεπτομέρειες για το πώς τυποποιήθηκαν οι παρεμβάσεις). <sup>43,44</sup> Συστήνουμε στους αναγνώστες να συμβουλευούνται τις δηλώσεις για τις μη φαρμακολογικές και τις βοτανικές παρεμβάσεις, ανάλογα με την περίπτωση.

### 6.3.6. Στοιχείο 6α. Ολοκληρωμένος προσδιορισμός των προκαθορισμένων κύριων και δευτερευόντων μέτρων έκβασης, περιλαμβανομένου του πώς και πότε αξιολογήθηκαν

*Παράδειγμα:* «Το πρωτεύον καταληκτικό σημείο με σεβασμό στην αποτελεσματικότητα στην ψωρίαση ήταν το ποσοστό των ασθενών που επιτυγχάνουν μια βελτίωση 75% στη δραστηριότητα της ψωρίασης από την αρχή έως τις 12 εβδομάδες, όπως μετρήθηκε με τον PASI (δείκτη

περιοχής και σοβαρότητας της ψωρίασης). Πρόσθετες αναλύσεις πραγματοποιήθηκαν στην αλλαγή του ποσοστού των αποτελεσμάτων PASI και στη βελτίωση των βλαβών της ψωρίασης». <sup>103</sup>

*Εξήγηση:* Όλες οι ΤΚΔ αξιολογούν τις μεταβλητές ανταπόκρισης, ή τις εκβάσεις (καταληκτικά σημεία), για τις οποίες συγκρίνονται οι ομάδες. Οι περισσότερες δοκιμές έχουν αρκετές εκβάσεις, μερικές από τις οποίες είναι περισσότερο ενδιαφέρουσες από άλλες. Η μέτρηση της κύριας έκβασης είναι η προκαθορισμένη έκβαση που θεωρείται μέγιστης σπουδαιότητας από τους σχετικούς ενδιαφερόμενους (όπως ασθενείς, φορείς χάραξης πολιτικής, κλινικούς ιατρούς, χρηματοδότες) και είναι συνήθως αυτή που χρησιμοποιείται στον υπολογισμό του μεγέθους του δείγματος (βλέπε στοιχείο 7). Μερικές δοκιμές μπορεί να έχουν περισσότερες από μία κύριες εκβάσεις. Έχοντας αρκετές κύριες εκβάσεις, εν τούτοις, υφίστανται τα προβλήματα ερμηνείας που συνδέονται με την πολλαπλότητα των αναλύσεων (βλέπε στοιχεία 18 και 20) και δεν συστήνονται. Οι κύριες εκβάσεις πρέπει να είναι σαφώς καθορισμένες, όπως στην αναφορά μιας ΤΚΔ. Άλλες ενδιαφέρουσες εκβάσεις είναι οι δευτερεύουσες εκβάσεις (πρόσθετες εκβάσεις). Μπορεί να υπάρχουν αρκετές δευτερεύουσες εκβάσεις, οι οποίες περιλαμβάνουν συχνά τα απρόβλεπτα ή τα ακούσια αποτελέσματα της παρέμβασης (βλέπε στοιχείο 19), αν και οι βλαβερές συνέπειες πρέπει πάντα να αντιμετωπίζονται ως σημαντικές, ανεξάρτητα με το γεγονός αν αποκαλούνται κύριες ή δευτερεύουσες.

Όλες οι μετρήσεις έκβασης, είτε αρχικές είτε δευτερεύουσες, πρέπει να προσδιοριστούν και να καθοριστούν πλήρως. Η αρχή εδώ αναφέρεται στο γεγονός ότι οι παρεχόμενες πληροφορίες πρέπει να είναι επαρκείς για να επιτρέπουν σε άλλους να χρησιμοποιούν τις ίδιες εκβάσεις. <sup>102</sup> Όταν αξιολογούνται οι εκβάσεις σε διάφορα χρονικά σημεία μετά από την τυχαιοποίηση, οι συγγραφείς πρέπει επίσης να δείξουν το προκαθορισμένο χρονικό σημείο του κύριου ενδιαφέροντος. Για πολλές, μη φαρμακολογικές παρεμβάσεις είναι χρήσιμο να διευκρινιστεί ποιος αξιολόγησε τις εκβάσεις (για παράδειγμα, εάν απαιτούνται πρόσθετες δεξιότητες για να το κάνεις) και πόσοι αξιολογητές υπήρχαν. <sup>43</sup>

Όπου είναι διαθέσιμη και κατάλληλη, θα πρέπει να αναφερθεί η χρήση των ήδη αναπτυγμένων και επικυρωμένων κλιμάκων ή των οδηγιών συναίνεσης, <sup>104,105</sup> για να βελτιώσουν την ποιότητα της μέτρησης και για να συνδράμουν στη σύγκριση με παρόμοιες μελέτες. <sup>106</sup> Για παράδειγμα, η αξιολόγηση της ποιότητας ζωής είναι πιθανό να βελτιωθεί με τη χρησιμοποίηση επικυρωμένων οργάνων. <sup>107</sup> Οι συγγραφείς πρέπει να επισημάνουν την προέλευση και τις ιδιότητες των κλιμάκων.



Περισσότερες από 70 εκβάσεις χρησιμοποιήθηκαν σε 196 ΤΚΔ των μη στεροειδών αντιφλεγμονωδών φαρμάκων για τη ρευματοειδή αρθρίτιδα<sup>108</sup> και 640 διαφορετικά όργανα έχουν χρησιμοποιηθεί σε 2.000 δοκιμές στη σχιζοφρένεια, από τα οποία 369 χρησιμοποιήθηκαν μόνο μία φορά.<sup>33</sup> Η έρευνα σε 149 από τις 2.000 δοκιμές έδειξε ότι οι αδημοσίευτες κλίμακες ήταν μια πηγή σφάλματος. Στις μη φαρμακολογικές δοκιμές, το ένα τρίτο των ισχυρισμών της ανωτερότητας της θεραπείας που βασίστηκε στις αδημοσίευτες κλίμακες δεν θα είχαν γίνει εάν είχε χρησιμοποιηθεί μια δημοσιευμένη κλίμακα.<sup>109</sup> Παρόμοια δεδομένα έχουν αναφερθεί αλλού.<sup>110,111</sup> Μόνο το 45% μιας κούρτης από 519 ΤΚΔ που δημοσιεύτηκαν το 2000 διευκρίνισε την κύρια έκβαση.<sup>16</sup> Αυτό συγκρίνεται με το 53% για μια παρόμοια κούρτη 614 ΤΚΔ που δημοσιεύτηκαν το 2006.<sup>17</sup>

### 6.3.7. Στοιχείο 6β. Οποιοσδήποτε αλλαγές στις εκβάσεις της δοκιμής ύστερα από την έναρξη της δοκιμής, με αιτίες

*Παράδειγμα:* «Το αρχικό πρωτεύον καταληκτικό σημείο ήταν η θνησιμότητα από όλες τις αιτίες, αλλά κατά τη διάρκεια μιας καλυμμένης (masked) ανάλυσης η επιτροπή ασφάλειας και επιτήρησης δεδομένων (data safety and monitoring board) σημείωσε ότι η συνολική θνησιμότητα ήταν χαμηλότερη απ' ό,τι είχε προβλεφθεί και ότι η μελέτη δεν θα μπορούσε να ολοκληρωθεί με το μέγεθος του δείγματος και την ισχύ που σχεδιάστηκε αρχικά. Η οργανωτική επιτροπή, επομένως, αποφάσισε να υιοθετήσει και τα άλλα κύρια καταληκτικά σημεία της θνησιμότητας από όλες τις αιτίες (το αρχικό πρωτεύον καταληκτικό σημείο), σε συνδυασμό με τη θνησιμότητα από όλες τις αιτίες ή τις εισαγωγές σε νοσοκομείο λόγω καρδιαγγειακών συμβαμάτων (το πρώτο προκαθορισμένο δευτερεύον καταληκτικό σημείο)».<sup>112</sup>

*Εξήγηση:* Υπάρχουν πολλοί λόγοι για αποκλίσεις από το αρχικό πρωτόκολλο της μελέτης (βλέπε στοιχείο 24). Οι συγγραφείς πρέπει να αναφέρουν όλες τις σημαντικές μεταβολές στο πρωτόκολλο, περιλαμβανομένων των μη σχεδιασμένων αλλαγών στα κριτήρια καταλληλότητας, τις παρεμβάσεις, τις εξετάσεις, τη συλλογή δεδομένων, τις μεθόδους ανάλυσης και τις εκβάσεις. Τέτοιες πληροφορίες δεν αναφέρονται πάντα.

Όπως επισημάνθηκε προηγουμένως (βλέπε στοιχείο 6α), οι περισσότερες δοκιμές καταγράφουν πολλαπλές εκβάσεις, με τον κίνδυνο ότι τα αποτελέσματα θα καταγραφούν για ένα μόνο επιλεγμένο υποσύνολο (βλέπε στοιχείο 17). Ο προκαθορισμός και η καταγραφή των κύριων και των δευτερευουσών εκβάσεων (βλέπε στοιχείο 6α) πρέπει να εξαλείφουν έναν τέτοιο κίνδυνο. Σε μερικές δοκιμές, εν τούτοις, οι περιστάσεις απαιτούν μια αλλαγή στον τρόπο

με τον οποίο αξιολογείται μια έκβαση ή ακόμη και, όπως στο παραπάνω παράδειγμα, μια μεταβολή προς μια διαφορετική έκβαση. Για παράδειγμα, μπορεί να υπάρξουν εξωτερικές ενδείξεις από άλλες δοκιμές ή συστηματικές ανασκοπήσεις που προτείνουν ότι το καταληκτικό σημείο μπορεί να μην είναι κατάλληλο ή η στρατολόγηση ή το συνολικό ποσοστό του συμβάντος στη δοκιμή μπορεί να είναι χαμηλότερο από το αναμενόμενο.<sup>112</sup> Η τροποποίηση ενός καταληκτικού σημείου βασισμένου σε μη τυφλοποιημένα δεδομένα είναι περισσότερο προβληματική, αν και μπορεί να διευκρινιστεί στο πλαίσιο ενός προσαρμοστικού σχεδιασμού της δοκιμής.<sup>88</sup> Οι συγγραφείς πρέπει να αναγνωρίσουν και να εξηγήσουν οποιοσδήποτε τέτοιες αλλαγές. Παρομοίως, οποιοσδήποτε μεταβολές μετά από την έναρξη της δοκιμής στο σχεδιασμό των εκβάσεων όπως κύριες ή δευτερεύουσες θα πρέπει να αναφερθούν και να επεξηγηθούν.

Από μια σύγκριση των πρωτοκόλλων και των δημοσιεύσεων 102 τυχαιοποιημένων δοκιμών φάνηκε ότι το 62% των δημοσιευμένων δοκιμών είχαν τουλάχιστον μία κύρια έκβαση που άλλαξε, εισήχθη ή παραλείφθηκε σε σχέση με το πρωτόκολλο.<sup>55</sup> Οι κύριες εκβάσεις επίσης διέφεραν μεταξύ των πρωτοκόλλων και των δημοσιεύσεων στο 40% μιας κούρτης 48 δοκιμών που χρηματοδοτήθηκαν από τα Καναδικά Ιδρύματα Υγείας και Έρευνας.<sup>113</sup> Καμιά από τις προαναφερθείσες 150 δοκιμές δεν είχε αλλαγές από το πρωτόκολλο με κάποια εξήγηση. Παρόμοια αποτελέσματα από άλλες μελέτες έχουν καταγραφεί πρόσφατα σε μια συστηματική ανασκόπηση των εμπειρικών μελετών που εξετάζουν το σφάλμα καταγραφής των εκβάσεων.<sup>57</sup>

### 6.3.8. Στοιχείο 7α. Πώς καθορίστηκε το μέγεθος του δείγματος;

*Παράδειγμα:* «Για να ανιχνευτεί μια μείωση της μετεγχειρητικής παραμονής στο νοσοκομείο (ΜΠΝ) των 3 ημερών (SD, τυπική απόκλιση 5 ημέρες), η οποία είναι σε συμφωνία με τη μελέτη των Lobo et al,<sup>17</sup> με ένα αμφίπλευρο επίπεδο σημαντικότητας 5% και μια ισχύ 80%, ήταν αναγκαίο ένα μέγεθος δείγματος 50 ασθενών ανά ομάδα, λαμβάνοντας υπ' όψη ένα προβλεπόμενο ποσοστό απώλειας 10%. Για να στρατολογηθεί αυτός ο αριθμός των ασθενών προβλεπόταν μια 12μηνη περίοδος εισόδου».<sup>114</sup>

«Βασισμένος σε μια αναμενόμενη επίπτωση του κύριου σύνθετου καταληκτικού σημείου 11% στα 2,25 έτη στην ομάδα εικονικού φαρμάκου, υπολογίσαμε ότι θα χρειαζόμασταν 950 συμβάντα κύριου καταληκτικού σημείου και ένα μέγεθος δείγματος 9.650 ασθενών για να υπάρχει 90% ισχύς, ώστε να ανιχνεύσουμε μια σημαντική διαφορά μεταξύ του ivabradine και του εικονικού φαρμάκου, που



αντιστοιχεί σε μια μείωση 19% του σχετικού κινδύνου (με ένα αμφίπλευρο σφάλμα τύπου 1, 5%). Αρχικά, σχεδιάσαμε μια οδηγούμενη από τα συμβάντα δοκιμή, και προγραμματίστηκε να διακοπεί όταν σημειώθηκαν 950 συμβάντα κύριου καταληκτικού σημείου. Εν τούτοις, η επίπτωση του κύριου καταληκτικού σημείου ήταν υψηλότερη απ' ό,τι προβλεπόταν, ίσως λόγω των βασικών χαρακτηριστικών των στρατολογημένων ασθενών, οι οποίοι είχαν υψηλότερο κίνδυνο απ' ό,τι αναμενόταν (π.χ. χαμηλότερο ποσοστό σταδίου I κατά NYHA και υψηλότερα ποσοστά διαβήτη και αρτηριακής υπέρτασης). Υπολογίσαμε ότι όταν σημειώθηκαν 950 συμβάντα κύριου καταληκτικού σημείου, οι πιο πρόσφατα ενταγμένοι ασθενείς θα είχαν αντιμετωπιστεί μόνο για περίπου 3 μήνες. Επομένως, τον Ιανουάριο του 2007, η εκτελεστική επιτροπή αποφάσισε να αλλάξει τη μελέτη από οδηγούμενη από τα συμβάντα σε οδηγούμενη από το χρόνο, και να τη συνεχίσει έως ότου οι ασθενείς που είχαν τυχαιοποιηθεί τελευταίοι να έχουν παρακολουθηθεί για 12 μήνες. Αυτή η τροποποίηση δεν άλλαξε την προγραμματισμένη διάρκεια μελέτης των 3 ετών». <sup>115</sup>

*Εξήγηση:* Για επιστημονικούς και ηθικούς λόγους, το μέγεθος του δείγματος μιας δοκιμής πρέπει να σχεδιαστεί προσεκτικά, με μια ισορροπία μεταξύ των ιατρικών και των στατιστικών εκτιμήσεων. Ιδανικά, μια μελέτη πρέπει να είναι αρκετά μεγάλη, να έχει μια υψηλή πιθανότητα (ισχύ) ανίχνευσης ως στατιστικά σημαντικής μιας κλινικά σημαντικής διαφοράς ενός δεδομένου μεγέθους, εάν υπάρχει μια τέτοια διαφορά. Το μέγεθος του αποτελέσματος που θεωρείται σημαντικό συσχετίζεται αντίστροφα με το μέγεθος του δείγματος που είναι απαραίτητο για να ανιχνευτεί δηλαδή, τα μεγάλα δείγματα είναι αναγκαία για να ανιχνευτούν μικρές διαφορές. Στοιχεία του υπολογισμού του μεγέθους του δείγματος είναι (α) οι εκτιμώμενες εκβάσεις σε κάθε ομάδα (που υπονοεί τις κλινικά σημαντικές διαφορές στόχων μεταξύ των ομάδων παρέμβασης), (β) το επίπεδο σημαντικότητας α (σφάλμα τύπου 1), (γ) η στατιστική ισχύς (το β, το επίπεδο σφάλματος [τύπου 2]) και (δ) για συνεχείς εκβάσεις, η τυπική απόκλιση των μετρήσεων. <sup>116</sup> Η αλληλεπίδραση αυτών των στοιχείων και η καταγραφή τους θα διαφέρει για δοκιμές συρροής <sup>40</sup> και μη υπεροχής, καθώς και για δοκιμές ισοδυναμίας. <sup>39</sup>

Οι συγγραφείς πρέπει να δείξουν πώς καθορίστηκε το μέγεθος του δείγματος. Εάν χρησιμοποιήθηκε ένας επίσημος υπολογισμός ισχύος, οι συγγραφείς πρέπει να προσδιορίσουν την αρχική έκβαση στην οποία ήταν βασισμένος ο υπολογισμός (βλέπε στοιχείο βα), όλα τα μεγέθη που χρησιμοποιήθηκαν στον υπολογισμό, καθώς και το μέγεθος δείγματος-στόχου που προκύπτει ανά ομάδα μελέτης. Προτιμότερο είναι να αναφερθεί το αναμενόμενο

αποτέλεσμα στην ομάδα ελέγχου και η διαφορά μεταξύ των ομάδων που δεν θα πρέπει να παραβλεφθεί. Εναλλακτικά, οι συγγραφείς θα μπορούσαν να παρουσιάσουν το ποσοστό με το συμβάν ή τη μέση τιμή για κάθε ομάδα που χρησιμοποιείται στους υπολογισμούς τους. Λεπτομέρειες θα πρέπει να δοθούν για οποιαδήποτε ανοχή σχετικά με την απώλεια ή τη μη συμμόρφωση κατά τη διάρκεια της μελέτης.

Μερικοί μεθοδολόγοι έχουν γράψει ότι οι αποκαλούμενες μικρής ισχύος δοκιμές μπορεί να είναι αποδεκτές, επειδή μπόρεσαν τελικά να συνδυαστούν σε μια συστηματική ανασκόπηση και μετα-ανάλυση <sup>117-119</sup> και επειδή κάποιες πληροφορίες είναι καλύτερες από καμιά πληροφορία. Πρέπει να σημειωθεί ότι υπάρχουν σημαντικοί περιοριστικοί όροι, όπως ότι η δοκιμή πρέπει να είναι αμερόληπτη, κατάλληλα αναφερόμενη, και δημοσιευμένη ανεξάρτητα από τα αποτελέσματα, επειδή με αυτόν τον τρόπο γίνεται διαθέσιμη για μετα-ανάλυση. <sup>118</sup> Από την άλλη πλευρά, πολλοί ιατρικοί ερευνητές ανησυχούν ότι οι μικρής ισχύος δοκιμές με απροσδιόριστα αποτελέσματα θα παραμείνουν αδημοσίετες και επιμένουν ότι όλες οι δοκιμές πρέπει ξεχωριστά να έχουν «επαρκή ισχύ». Αυτή η συζήτηση θα συνεχιστεί και τα μέλη της Ομάδας του CONSORT έχουν ποικίλες απόψεις. Αυστηρά εν τούτοις, η συζήτηση και εκείνες οι απόψεις είναι ασήμαντες στην αναφορά μιας δοκιμής. Οποιαδήποτε και αν είναι η ισχύς μιας δοκιμής, οι συγγραφείς πρέπει να αναφέρουν κατάλληλα το προοριζόμενο μέγεθός της με όλες τις μεθόδους και τις υποθέσεις. <sup>118</sup> Αυτό αποκαλύπτει διαφανώς την ισχύ της δοκιμής στους αναγνώστες και τους παρέχει ένα μέτρο με το οποίο αξιολογούν εάν η δοκιμή πέτυχε το προγραμματισμένο μέγεθός της.

Σε μερικές δοκιμές, οι ενδιάμεσες αναλύσεις χρησιμοποιούνται για να βοηθήσουν να αποφασιστεί εάν πρόκειται να διακοπούν νωρίς ή να συνεχίσουν να στρατολογούν μερικές φορές πέρα από το προγραμματισμένο τέλος της δοκιμής (βλέπε στοιχείο 7β). Εάν το πραγματικό μέγεθος του δείγματος διαφέρει από το αρχικά προοριζόμενο δείγμα για κάποιον άλλο λόγο (για παράδειγμα, λόγω της κακής στρατολόγησης ή της αναθεώρησης του αρχικά υπολογισμένου μεγέθους του δείγματος [target sample size]), πρέπει να δοθεί εξήγηση.

Οι αναφορές των μελετών με μικρά δείγματα εμπειρικών συχνά το λανθασμένο συμπέρασμα ότι οι ομάδες παρέμβασης δεν διαφέρουν, όταν στην πραγματικότητα μελετήθηκαν πολύ λίγοι ασθενείς για να κάνουν έναν τέτοιο ισχυρισμό. <sup>120</sup> Οι ανασκοπήσεις των δημοσιευμένων δοκιμών έχουν, με συνέπεια, διαπιστώσει ότι ένα μεγάλο τμήμα των δοκιμών έχει χαμηλή ισχύ στο να ανιχνεύσει

κλινικά σημαντικά θεραπευτικά αποτελέσματα.<sup>121–123</sup> Στην πραγματικότητα, μικρές αλλά κλινικώς σημαντικά αληθινές διαφορές είναι πιθανότερο να υπάρχουν από τις μεγάλες διαφορές, αλλά απαιτούνται μεγάλες δοκιμές για να ανιχνευτούν.<sup>124</sup>

Γενικά, τα αναφερόμενα μεγέθη των δειγμάτων στις δοκιμές φαίνονται μικρά. Το διάμεσο μέγεθος του δείγματος ήταν 54 ασθενείς σε 196 δοκιμές για την αρθρίτιδα,<sup>108</sup> 46 ασθενείς σε 73 δοκιμές στη Δερματολογία,<sup>8</sup> και 65 ασθενείς σε 2.000 δοκιμές για τη σχιζοφρένεια.<sup>33</sup> Τα μικρά μεγέθη των δειγμάτων είναι σύμφωνα με μια μελέτη 519 δοκιμών που καταχωρήθηκαν στο PubMed το Δεκέμβριο του 2000<sup>16</sup> και παρόμοια κούρτη δοκιμών (n=616) που καταχωρήθηκαν στο PubMed το 2006,<sup>17</sup> όπου ο διάμεσος αριθμός των ασθενών που στρατολογήθηκαν για τις παράλληλες ομάδες δοκιμών ήταν 80 και τα δύο έτη. Επί πλέον, πολλές ανασκοπήσεις έχουν διαπιστώσει ότι λίγοι συγγραφείς αναφέρουν πώς καθόρισαν το μέγεθος του δείγματος.<sup>8,14,32,33,123</sup>

Υπάρχει λίγη αξία στον εκ των υστέρων υπολογισμό της στατιστικής ισχύος, χρησιμοποιώντας τα αποτελέσματα μιας δοκιμής· αντ' αυτού, η ισχύς είναι κατάλληλα υποδεικνύομενη από τα διαστήματα εμπιστοσύνης (βλέπε στοιχείο 17).<sup>125</sup>

### 6.3.9. Στοιχείο 7β. Όπου είναι εφαρμόσιμο, εξήγηση των ενδιάμεσων αναλύσεων και των οδηγιών διακοπής

*Παραδείγματα:* «Δύο ενδιάμεσες αναλύσεις πραγματοποιήθηκαν κατά τη διάρκεια της δοκιμής. Τα επίπεδα σημαντικότητας διατήρησαν μια γενική τιμή  $p$  που αντιστοιχούσε σε 0,05 και υπολογίστηκαν σύμφωνα με τα όρια διακοπής των O'Brien-Fleming. Η εν λόγω τελική ανάλυση χρησιμοποίησε μια τιμή  $Z$  1,985 με μια σχετιζόμενη τιμή  $p=0,0471$ ».<sup>126</sup>

«Ένα ανεξάρτητο συμβούλιο επιτήρησης δεδομένων και ασφάλειας ανασκόπησε περιοδικά τα στοιχεία αποτελεσματικότητας και ασφάλειας. Οι κανόνες διακοπής βασίστηκαν στα τροποποιημένα όρια των Haybittle-Peto των 4 τυπικών αποκλίσεων (SD) στο πρώτο μισό της μελέτης και 3 SD στο δεύτερο μισό για τα δεδομένα αποτελεσματικότητας, και των 3 SD στο πρώτο μισό της μελέτης και 2 SD στο δεύτερο μισό για τα δεδομένα ασφάλειας. Δύο επίσημες ενδιάμεσες αναλύσεις της αποτελεσματικότητας πραγματοποιήθηκαν, όταν το 50% και το 75% του αναμενόμενου αριθμού των αρχικών συμβάντων είχε συλλεχθεί. Δεν πραγματοποιήθηκε κάποια διόρθωση της αναφερόμενης τιμής  $p$  για τις συγκεκριμένες ενδιάμεσες δοκιμές».<sup>127</sup>

*Εξήγηση:* Πολλές δοκιμές στρατολογούν συμμετέχο-

ντες κατά τη διάρκεια μιας μεγάλης περιόδου. Εάν μια παρέμβαση λειτουργεί ιδιαίτερα καλά ή άσχημα, η μελέτη μπορεί να απαιτηθεί να τερματιστεί νωρίς για ηθικούς λόγους. Αυτή η ανησυχία μπορεί να αντιμετωπιστεί με την εξέταση των αποτελεσμάτων όσο συσσωρεύονται τα δεδομένα, κατά προτίμηση από μια ανεξάρτητη επιτροπή επιτήρησης δεδομένων. Εν τούτοις, η πραγματοποίηση πολλαπλών στατιστικών ελέγχων των συσσωρευμένων δεδομένων χωρίς κατάλληλη διόρθωση μπορεί να οδηγήσει σε λανθασμένα αποτελέσματα και ερμηνείες.<sup>128</sup> Αν τα συσσωρευμένα δεδομένα από μια δοκιμή εξετάστηκαν σε πέντε ενδιάμεσες αναλύσεις που χρησιμοποιούν μια τιμή  $p=0,05$ , το συνολικό ψευδώς θετικό ποσοστό είναι πλησιέστερα στο 19% απ' ό,τι στο συμβατικό 5%.

Αρκετές ομάδες διαδοχικών στατιστικών μεθόδων είναι διαθέσιμες για να σταθμίσουν για πολλαπλές αναλύσεις,<sup>129–131</sup> και η χρήση τους πρέπει να προκαθοριστεί στο πρωτόκολλο της δοκιμής. Με αυτές τις μεθόδους, τα δεδομένα συγκρίνονται σε κάθε ενδιάμεση ανάλυση, ενώ μια τιμή  $p$  μικρότερη από την κριτική τιμή που καθορίζεται με τη μέθοδο των διαδοχικών συγκρίσεων δείχνει τη στατιστική σημαντικότητα. Μερικοί διεξάγοντες τις δοκιμές χρησιμοποιούν τις μεθόδους των διαδοχικών ομάδων ως ενίσχυση στη λήψη απόφασης,<sup>132</sup> ενώ άλλοι τις μεταχειρίζονται ως επίσημο κανόνα διακοπής (με την πρόθεση ότι η δοκιμή θα διακοπεί εάν η παρατηρηθείσα τιμή  $p$  είναι μικρότερη από την κριτική τιμή).

Οι συγγραφείς πρέπει να αναφέρουν αν αυτοί ή μια επιτροπή επιτήρησης δεδομένων έριξαν πολλαπλές «ματιές» στα δεδομένα και, σε αυτή την περίπτωση, πόσες υπήρξαν, τι τις προκάλεσε, τις στατιστικές μεθόδους που εφαρμόστηκαν (περιλαμβανομένου οποιουδήποτε επίσημου κανόνα διακοπής), και εάν προγραμματίστηκαν πριν από την έναρξη της δοκιμής, πριν η επιτροπή επιτήρησης δεδομένων δει οποιαδήποτε ενδιάμεσα δεδομένα από τη διανομή θεραπείας, ή κάποια στιγμή έκτοτε. Αυτές οι πληροφορίες συχνά δεν περιλαμβάνονται στις δημοσιευμένες εκθέσεις των δοκιμών,<sup>133</sup> ακόμη και στις δοκιμές που αναφέρουν διακοπή νωρίτερα από το σχεδιασμένο.<sup>134</sup>

### 6.3.10. Στοιχείο 8α. Μέθοδος που εφαρμόστηκε για τη δημιουργία της ακολουθίας τυχαίας διανομής

*Παραδείγματα:* «Ανεξάρτητοι φαρμακοποιοί διένειμαν είτε ενεργείς είτε εικονικές αναπνευστικές συσκευές, σύμφωνα με έναν κατάλογο τυχαίοποίησης που παρήχθη από υπολογιστή».<sup>63</sup>

«Για τη διανομή των συμμετεχόντων, χρησιμοποιήθηκε μια λίστα τυχαίων αριθμών παραγόμενη από υπολογιστή».<sup>135</sup>

**Εξήγηση:** Οι συμμετέχοντες πρέπει να διανεμηθούν στις ομάδες σύγκρισης της δοκιμής με βάση μια διαδικασία πιθανοτήτων (τυχαία) που χαρακτηρίζεται από μη προβλεψιμότητα (βλέπε πλαίσιο 1). Οι συγγραφείς πρέπει να παρέχουν επαρκείς πληροφορίες, με τις οποίες ο αναγνώστης να είναι σε θέση να αξιολογήσει τις μεθόδους που χρησιμοποιήθηκαν για να παραχθεί η τυχαία ακολουθία διανομής και την πιθανότητα σφάλματος στην ταξινόμηση σε ομάδες. Είναι σημαντικό ότι οι πληροφορίες για τη διαδικασία της τυχαιοποίησης πρέπει να περιλαμβάνονται στο σώμα του κυρίως άρθρου και όχι ως ξεχωριστό συμπληρωματικό αρχείο, το οποίο μπορεί να απωλέσει ο αναγνώστης.

Ο όρος «τυχαίος» έχει μια ακριβή τεχνική έννοια. Με την τυχαιοποιημένη διανομή, κάθε συμμετέχων έχει μια γνωστή πιθανότητα της λήψης κάθε παρέμβασης προτού ταξινομηθεί, αλλά η ταξινομημένη παρέμβαση καθορίζεται με μια διαδικασία τύχης και δεν μπορεί να προβλεφθεί. Εν τούτοις, ο όρος «τυχαίος» χρησιμοποιείται συχνά ακατάλληλα στη βιβλιογραφία για να περιγράψει δοκιμές στις οποίες χρησιμοποιήθηκαν μη τυχαίες, προβλέψιμες μέθοδοι διανομής, όπως η εναλλαγή, οι αριθμοί των νοσοκομείων ή η ημερομηνία γέννησης. Όταν οι ερευνητές χρησιμοποιούν τέτοιες μη τυχαίες μεθόδους, πρέπει να τις περιγράφουν επακριβώς και δεν πρέπει να χρησιμοποιούν τον όρο «τυχαίος» ή οποιαδήποτε παραλλαγή του. Ακόμη και ο όρος «ημι-τυχαίος» δεν είναι αποδεκτός για την περιγραφή τέτοιων δοκιμών. Δοκιμές βασισμένες σε μη τυχαίες μεθόδους γενικά αποφέρουν εσφαλμένα αποτελέσματα.<sup>2-4,136</sup> Το σφάλμα προκύπτει πιθανόν από την ανικανότητα να αποκρύψουν αυτά τα συστήματα διανομής επαρκώς (βλέπε στοιχείο 9).

Πολλές μέθοδοι παραγωγής ακολουθίας είναι επαρκείς. Εν τούτοις, οι αναγνώστες δεν μπορούν να κρίνουν την επάρκεια από τέτοιους όρους, όπως «τυχαιοποιημένη διανομή», «τυχαιοποίηση» ή «τυχαίος», χωρίς περαιτέρω επεξεργασία. Οι συγγραφείς πρέπει να διευκρινίζουν τη μέθοδο παραγωγής της ακολουθίας, όπως έναν πίνακα τυχαίων αριθμών ή μια γεννήτρια τυχαίων αριθμών από υπολογιστή. Η ακολουθία μπορεί να παραχθεί με τη διαδικασία της ελαχιστοποίησης (minimization), μια μη τυχαία, αλλά γενικά αποδεκτή μέθοδο (βλέπε πλαίσιο 2).

Σε μερικές δοκιμές, οι συμμετέχοντες διανεμήθηκαν σκόπιμα σε άνισους αριθμούς σε κάθε παρέμβαση: για παράδειγμα, προκειμένου να αποκτηθεί περισσότερη εμπειρία με μια νέα διαδικασία ή για να περιοριστεί το κόστος της δοκιμής. Σε τέτοιες περιπτώσεις, οι συγγραφείς πρέπει να εκθέσουν την αναλογία τυχαιοποίησης (για παράδειγμα, 2:1 ή δύο συμμετέχοντες στη θεραπεία ανά κάθε συμμετέχοντα ελέγχου) (βλέπε στοιχείο 3α).

Σε ένα αντιπροσωπευτικό δείγμα δοκιμών καταχωρημένων στο PubMed το 2000, μόνο το 21% ανέφερε μια επαρκή προσέγγιση στην παραγωγή τυχαιοποιημένης ακολουθίας.<sup>16</sup> Αυτό αυξήθηκε σε 34% για μια παρόμοια κοόρτη καταχωρημένων στο PubMed δοκιμών το 2006.<sup>17</sup> Σε περισσότερες από το 90% αυτών των περιπτώσεων, οι ερευνητές χρησιμοποίησαν ένα σύστημα παραγωγής τυχαίων αριθμών σε έναν υπολογιστή ή έναν τυχαίο πίνακα αριθμών.

### 6.3.11. Στοιχείο 8β. Τύπος τυχαιοποίησης: Λεπτομέρειες τυχόν περιορισμών (όπως χρήση μπλοκ [blocking] και μέγεθος μπλοκ)

**Παραδείγματα:** «Η ακολουθία τυχαιοποίησης δημιουργήθηκε χρησιμοποιώντας το στατιστικό λογισμικό Stata 9.0 (StataCorp, College Station, TX) και διαστρωματώθηκε ανά κέντρο με μια 1:1 διανομή, χρησιμοποιώντας τυχαιοποιημένα μεγέθη μπλοκ 2, 4 και 6».<sup>137</sup>

«Οι συμμετέχοντες ταξινομήθηκαν τυχαιοποιημένα, ακολουθώντας απλές διαδικασίες τυχαιοποίησης (τυχαίοι αριθμοί από υπολογιστή) σε μία από τις δύο ομάδες θεραπείας».<sup>138</sup>

**Εξήγηση:** Στις δοκιμές αρκετών εκατοντάδων συμμετεχόντων ή περισσότερων, η απλή τυχαιοποίηση μπορεί συνήθως να χρησιμοποιηθεί για να παραχθούν παρόμοιοι αριθμοί συμμετεχόντων στις δύο ομάδες της δοκιμής<sup>139</sup> και για να παραχθούν ομάδες που είναι κατά προσέγγιση συγκρίσιμες, από την άποψη γνωστών και άγνωστων προγνωστικών μεταβλητών.<sup>140</sup> Για τις μικρότερες δοκιμές (βλέπε στοιχείο 7α) –και ακόμη και για τις δοκιμές που δεν προορίζονται να είναι μικρές, καθώς μπορεί να διακοπούν πριν φθάσουν στο αρχικά υπολογισμένο μέγεθος (target size)– κάποια περιορισμένη τυχαιοποίηση (διαδικασίες για να βοηθήσουν να επιτύχει η ισορροπία μεταξύ των ομάδων στο μέγεθος ή στα χαρακτηριστικά) μπορεί να είναι χρήσιμη (βλέπε πλαίσιο 2).

Είναι σημαντικό να υποδειχθεί κατά πόσο δεν χρησιμοποιήθηκε κανένας περιορισμός, δηλώνοντάς το ή δηλώνοντας ότι εφαρμόστηκε «απλή τυχαιοποίηση». Διαιρητικά, οι μέθοδοι που χρησιμοποιήθηκαν για να περιορίσουν την τυχαιοποίηση, μαζί με τη μέθοδο που εφαρμόστηκε για την τυχαία επιλογή, πρέπει να καθοριστούν. Για την τυχαιοποίηση με μπλοκ, οι συγγραφείς πρέπει να παρέχουν λεπτομέρειες για τον τρόπο με τον οποίο παρήχθησαν τα μπλοκ (για παράδειγμα, χρησιμοποιώντας έναν εναλλασσόμενο σχεδιασμό μπλοκ με μια τυχαιοποιημένη διάταξη αριθμών από υπολογιστή), το μέγεθος ή τα μεγέθη των μπλοκ, και εάν το μέγεθος των μπλοκ ήταν σταθερό ή κυμαινόταν τυχαιοποιημένα. Επίσης,



οι πληροφορίες εκείνες που πρέπει να αναφερθούν είναι εάν οι διεξάγοντες τις δοκιμές ήταν ενήμεροι για το μέγεθος των μπλοκ, δεδομένου ότι τέτοια γνώση θα μπορούσε να οδηγήσει στο σπάσιμο του κώδικα. Οι συγγραφείς πρέπει να διευκρινίσουν εάν χρησιμοποιήθηκε διαστρωμάτωση, και σε αυτή την περίπτωση, ποιοι παράγοντες αναμίχθηκαν (όπως η περιοχή στρατολόγησης, το φύλο, το στάδιο της νόσου), οι τιμές κατηγοριοποίησης των στρωμάτων, καθώς και η μέθοδος που χρησιμοποιήθηκε για τον περιορισμό. Αν και η διαστρωμάτωση είναι μια χρήσιμη τεχνική, ειδικά για τις μικρότερες δοκιμές, είναι περίπλοκο να εφαρμοστεί και μπορεί να είναι αδύνατη εάν χρησιμοποιούνται πολλοί παράγοντες διαστρωμάτωσης. Εάν χρησιμοποιήθηκε η ελαχιστοποίηση (βλέπε πλαίσιο 2), έπρεπε να προσδιοριστεί σαφώς, όπως και οι μεταβλητές πρέπει να ενσωματωθούν μέσα στο σχήμα. Εάν χρησιμοποιήθηκε, πρέπει να υποδειχθεί ένα τυχαίο στοιχείο.

Μόνο το 9% των 206 δημοσιεύσεων των δοκιμών σε περιοδικά ειδικότητας<sup>23</sup> και το 39% 80 δοκιμών σε περιοδικά Γενικής Ιατρικής ανέφεραν χρήση της διαστρωμάτωσης.<sup>32</sup> Σε κάθε περίπτωση, μόνο οι μισές περίπου από τις δημοσιεύσεις ανέφεραν τη χρήση της περιορισμένης τυχαίοποίησης. Εν τούτοις, αυτές οι μελέτες και εκείνη των Adetugbo και Williams<sup>9</sup> διαπίστωσαν ότι τα μεγέθη των ομάδων θεραπείας σε πολλές δοκιμές ήταν τα ίδια ή παρόμοια, χωρίς να αναφερθεί η χρήση μπλοκ και διαστρωμάτωσης. Μια πιθανή εξήγηση για τη στενή ισορροπία των αριθμών είναι η πλημμελής αναφορά της χρήσης της περιορισμένης τυχαίοποίησης.

### 6.3.12. Στοιχείο 9. Μηχανισμός που χρησιμοποιήθηκε για την εφαρμογή της ακολουθίας τυχαίας διανομής (όπως διαδοχικά αριθμημένοι περιέκτες), περιγράφοντας οποιαδήποτε βήματα που τηρήθηκαν για την απόκρυψη της ακολουθίας μέχρι την εφαρμογή των παρεμβάσεων

*Παραδείγματα:* «Η δοξουκλίνη και το εικονικό φάρμακο ήταν σε μορφή καψουλών και ίδια στην εμφάνιση. Ήταν προσσκευασμένα σε φιαλίδια και διαδοχικά αριθμημένα για κάθε γυναίκα, σύμφωνα με το πρόγραμμα τυχαίοποίησης. Κάθε γυναίκα ταξινομήθηκε με έναν αριθμό και έλαβε τις κάψουλες στο αντίστοιχο προσσκευασμένο φιαλίδιο».<sup>146</sup>

«Η ακολουθία διανομής αποκρύφθηκε από τον ερευνητή (JR) που ενέγραφε και αξιολογούσε τους συμμετέχοντες στους διαδοχικά αριθμημένους, αδιαφανείς, σφραγισμένους και συρραμμένους φακέλους. Το φύλλο αλουμινίου μέσα στο φάκελο χρησιμοποιήθηκε για να καταστήσει το φάκελο στεγανό στο έντονο φως. Για να αποτραπεί η αποκάλυψη της ακολουθίας διανομής, στο φάκελο

αναγράφηκε το όνομα και η ημερομηνία γέννησης του συμμετέχοντα και φτιάχθηκε μια τηλεοπτική ταινία του σφραγισμένου φακέλου με ορατές τις λεπτομέρειες των συμμετεχόντων. Το καρμπόν μέσα στο φάκελο μετέφερε τις πληροφορίες επάνω στην κάρτα διανομής μέσα στο φάκελο και ένας δεύτερος ερευνητής (CC) αργότερα είδε τις τηλεοπτικές ταινίες για να εξασφαλίσει ότι οι φάκελοι ήταν ακόμη σφραγισμένοι, όταν αναγράφηκαν τα ονόματα των συμμετεχόντων σε αυτούς. Οι αντίστοιχοι φάκελοι ανοίχθηκαν μόνο αφού οι εγγεγραμμένοι συμμετέχοντες ολοκλήρωσαν όλες τις αρχικές βασικές μετρήσεις και ήταν η στιγμή να διανεμηθεί η παρέμβαση».<sup>147</sup>

*Εξήγηση:* Στο στοιχείο 8α συζητήθηκε η παραγωγή μιας απρόβλεπτης ακολουθίας ταξινομήσεων. Μεγάλης σπουδαιότητας είναι το πώς αυτή η ακολουθία εφαρμόζεται, όταν οι συμμετέχοντες εγγράφονται στη δοκιμή (βλέπε το πλαίσιο 1). Ένα πρόγραμμα διανομής που παρήχθη πρέπει να εφαρμοστεί, χρησιμοποιώντας απόκρυψη διανομής,<sup>23</sup> ένας κρίσιμος μηχανισμός που αποτρέπει την πρόγνωση της ταξινόμησης θεραπείας και έτσι προστατεύει εκείνους που εγγράφουν τους συμμετέχοντες από την επιρροή αυτής της γνώσης. Η απόφαση να γίνει αποδεκτός ή να απορριφθεί ένας συμμετέχων πρέπει να ληφθεί, και η συγκατάθεση μετά από ενημέρωση πρέπει να ληφθεί από το συμμετέχοντα, αγνοώντας την επόμενη ταξινόμηση στην ακολουθία.<sup>148</sup>

Η απόκρυψη διανομής δεν πρέπει να συγχέεται με την τυφλοποίηση (βλέπε στοιχείο 11). Η απόκρυψη διανομής επιδιώκει να αποτρέψει το σφάλμα επιλογής, προστατεύει την ακολουθία ταξινόμησης μέχρι τη διανομή και μπορεί πάντα να εφαρμοστεί επιτυχώς.<sup>2</sup> Αντίθετα, η τυφλοποίηση επιδιώκει να αποτρέψει το σφάλμα απόδοσης και διαπίστωσης, προστατεύει την ακολουθία μετά τη διανομή και δεν μπορεί να εφαρμοστεί πάντα.<sup>23</sup> Χωρίς επαρκή απόκρυψη διανομής, εν τούτοις, ακόμη και τυχαίες, μη προβλέψιμες ακολουθίες ταξινόμησης μπορεί να ανατραπούν.<sup>2,149</sup>

Η κεντρική ή από έναν «τρίτο φορέα» ανάθεση είναι ιδιαίτερα επιθυμητή. Πολλοί καλοί μηχανισμοί απόκρυψης διανομής ενσωματώνουν εξωτερική ανάμιξη. Η χρήση ενός φαρμακείου ή ενός κεντρικού συστήματος τηλεφωνικής τυχαίοποίησης είναι δύο συχνές τεχνικές. Τα αυτοματοποιημένα συστήματα ταξινόμησης είναι πιθανό να γίνονται πιο συχνά.<sup>150</sup> Όταν η εξωτερική ανάμιξη δεν είναι εφικτή, μια άριστη μέθοδος απόκρυψης διανομής είναι η χρήση των αριθμημένων περιεκτών. Οι παρεμβάσεις (συχνά φάρμακα) σφραγίζονται στους διαδοχικά αριθμημένους πανομοιότυπους περιέκτες, σύμφωνα με την ακολουθία διανομής.<sup>151</sup> Οι εσώκλειστες ταξινομήσεις σε διαδοχικά αριθμημένους, αδιαφανείς, σφραγισμένους φακέλους μπορεί



να είναι ένας καλός μηχανισμός απόκρυψης διανομής εάν αναπτύσσεται και επιτηρείται επιμελώς. Αυτή η μέθοδος μπορεί να αλλοιωθεί, εν τούτοις, ιδιαίτερα εάν είναι κακώς εκτελεσθείσα. Οι ερευνητές πρέπει να εξασφαλίσουν ότι οι φάκελοι είναι αδιαφανείς όταν κρατιούνται στο φως, και ανοίγονται διαδοχικά και μόνο αφού το όνομα του συμμετέχοντα, καθώς και άλλες λεπτομέρειες, γράφονται στον κατάλληλο φάκελο.<sup>143</sup>

Ένας αριθμός μεθοδολογικών μελετών παρέχει εμπειρικά δεδομένα για να υποστηρίξουν αυτά τα μέτρα προφύλαξης.<sup>152,153</sup> Δοκιμές στις οποίες η απόκρυψη της ακολουθίας διανομής ήταν ανεπαρκής ή ασαφής απέφεραν μεγαλύτερες εκτιμήσεις των θεραπευτικών αποτελεσμάτων απ' ό,τι οι δοκιμές στις οποίες οι συγγραφείς ανέφεραν επαρκή απόκρυψη διανομής. Αυτά τα ευρήματα παρέχουν ισχυρά εμπειρικά δεδομένα ότι η ανεπαρκής απόκρυψη διανομής συμβάλλει σε σφάλμα στον υπολογισμό των θεραπευτικών αποτελεσμάτων.

Παρά τη σπουδαιότητα του μηχανισμού της απόκρυψης διανομής, οι δημοσιεύσεις παραλείπουν συχνά τέτοιες λεπτομέρειες. Ο μηχανισμός που χρησιμοποιήθηκε για να διανείμει τις παρεμβάσεις παραλείφθηκε στο 89% των δημοσιεύσεων των δοκιμών για τη ρευματοειδή αρθρίτιδα,<sup>108</sup> στο 48% των δοκιμών στα περιοδικά Μαιευτικής και Γυναικολογίας<sup>23</sup> και στο 44% των δοκιμών σε περιοδικά γενικού ιατρικού ενδιαφέροντος.<sup>32</sup> Σε ένα ευρύτερα αντιπροσωπευτικό δείγμα όλων των τυχαιοποιημένων δοκιμών που καταχωρήθηκαν στο PubMed, μόνο στο 18% καταγράφηκε οποιοσδήποτε μηχανισμός απόκρυψης διανομής, αλλά κάποιες από εκείνες που κατέγραφαν μηχανισμούς ήταν ανεπαρκείς.<sup>16</sup>

**6.3.10. Στοιχείο 10. Ποιος δημιούργησε την ακολουθία τυχαίας διανομής, ποιος ενέγραψε τους συμμετέχοντες και ποιος ταξινόμησε τους συμμετέχοντες στις παρεμβάσεις;**

*Παραδείγματα:* «Προσδιορισμός του εάν ένας ασθενής θα θεραπευόταν με στρεπτομυκίνη και κατάκλιση (περίπτωση S) ή με κατάκλιση μόνο (περίπτωση C) έγινε με αναφορά σε μια στατιστική σειρά βασισμένη σε τυχαίους δειγματοληπτικούς αριθμούς που επιλέχθηκαν ανά φύλο και ανά κέντρο από τον καθηγητή Bradford Hill. Οι λεπτομέρειες της σειράς ήταν άγνωστες σε οποιονδήποτε από τους ερευνητές ή στο συντονιστή... Μετά την αποδοχή ενός ασθενούς από την επιτροπή, και πριν από την είσοδο στο κέντρο στρεπτομυκίνης, ο κατάλληλος αριθμημένος φάκελος ανοίχθηκε στο κεντρικό γραφείο. Η κάρτα στο εσωτερικό ανέφερε εάν ο ασθενής επρόκειτο να είναι μια περίπτωση S ή μια περίπτωση C, και αυτή η πληροφορία δόθηκε έπειτα στον ιατρικό υπάλληλο του κέντρου».<sup>24</sup>

«Οι λεπτομέρειες της διανεμημένης ομάδας δόθηκαν σε χρωματιστές κάρτες που περιλαμβάνονταν σε διαδοχικά αριθμημένους, αδιαφανείς, σφραγισμένους φακέλους. Αυτοί προετοιμάστηκαν στο NPEU και φυλάχθηκαν σε μια συμφωνηθείσα θέση σε κάθε πτέρυγα. Η τυχαιοποίηση πραγματοποιήθηκε στο τέλος του δεύτερου σταδίου του τοκετού, όταν η μαία θεώρησε ότι ήταν επικείμενος ένας κολπικός τοκετός. Για να εισαχθεί μια γυναίκα στη μελέτη, η μαία άνοιγε τον επόμενο διαδοχικά αριθμημένο φάκελο».<sup>154</sup>

«Η τυχαιοποίηση με μπλοκ ήταν προετοιμασμένη από έναν ερευνητή, χωρίς κλινική εμπλοκή στη δοκιμή χρησιμοποιώντας έναν κατάλογο τυχαίων αριθμών που παρήχθη με τη βοήθεια ενός υπολογιστή. Διαστρωματώσαμε, σύμφωνα με την εισαγωγή, για μια σχετική ογκολογική διαδικασία. Αφότου η νοσηλεύτρια της έρευνας είχε λάβει τη συγκατάθεση του ασθενούς, τηλεφώνησε σε μια επαφή που ήταν ανεξάρτητη από τη διαδικασία στρατολόγησης για την αποστολή της διανομής».<sup>155</sup>

*Εξήγηση:* Όπως σημειώνεται στο στοιχείο 9, η απόκρυψη της διανεμημένης παρέμβασης κατά τη διάρκεια της εγγραφής είναι ιδιαίτερα σημαντική. Κατά συνέπεια, εκτός από τη γνώση των χρησιμοποιούμενων μεθόδων, είναι επίσης σημαντικό να αντιληφθείς πώς εφαρμόστηκε η τυχαία ακολουθία. Συγκεκριμένα, ποιος παρήγαγε την ακολουθία διανομής, ποιος ενέγραψε τους συμμετέχοντες και ποιος ταξινόμησε τους συμμετέχοντες στις ομάδες της δοκιμής.

Η διαδικασία της τυχαιοποίησης των συμμετεχόντων σε μια δοκιμή έχει τρία διαφορετικά στάδια: Παραγωγή ακολουθίας, απόκρυψη διανομής και εφαρμογή (βλέπε πλαίσιο 3). Αν και οι ίδιοι άνθρωποι μπορούν να πραγματοποιήσουν περισσότερες από μία διαδικασίες κάτω από κάθε στάδιο, οι ερευνητές πρέπει να προσπαθούν για τον πλήρη διαχωρισμό των ανθρώπων οι οποίοι ασχολούνται με την παραγωγή και την απόκρυψη διανομής από τους ανθρώπους που ασχολούνται με την εφαρμογή των ταξινομήσεων. Κατά συνέπεια, εάν κάποιος εμπλέκεται στα στάδια παραγωγής της ακολουθίας ή της απόκρυψης διανομής, ιδανικά δεν πρέπει να συμμετέχει στο στάδιο της εφαρμογής.

Ακόμη και με άψογη παραγωγή ακολουθίας και απόκρυψη διανομής, η αποτυχία να διαχωριστεί η δημιουργία και η απόκρυψη της ακολουθίας διανομής από την ταξινόμηση σε ομάδα μελέτης μπορεί να εισάγει σφάλμα. Για παράδειγμα, το πρόσωπο που παρήγαγε μια ακολουθία διανομής θα μπορούσε να διατηρήσει ένα αντίγραφο και να το συμβουλευτεί, παίρνοντας συνέντευξη από τους πιθανούς συμμετέχοντες σε μια δοκιμή. Κατά συνέπεια, εκείνο το πρόσωπο θα μπορούσε να προκαλέσει σφάλμα

στην εγγραφή ή στη διαδικασία ταξινόμησης, ανεξάρτητα από τη μη προβλεψιμότητα της ακολουθίας ταξινόμησης. Οι ερευνητές πρέπει έπειτα να εξασφαλίσουν ότι το πρόγραμμα ταξινόμησης είναι απρόβλεπτο και κλειδωμένο μακριά (όπως σε ένα ασφαλές κιβώτιο κατάθεσης, σε ένα κτήριο μάλλον απρόσιτο από την τοποθεσία εγγραφής), ακόμη και από το πρόσωπο που το παρήγαγε. Η δημοσίευση της δοκιμής πρέπει να διευκρινίζει πού διατήρησαν οι ερευνητές τη λίστα διανομής.

#### 6.3.14. Στοιχείο 11α. Αν έγινε, ποιος ήταν τυφλοποιημένος μετά από την ταξινόμηση των παρεμβάσεων (για παράδειγμα, οι συμμετέχοντες, οι προμηθευτές φροντίδας, αυτοί που αξιολογούν τις εκβάσεις) και πώς εφαρμόστηκε

*Παραδείγματα:* «Ενώ οι ασθενείς και οι ιατροί που διανεμήθηκαν σε ομάδα παρέμβασης γνώριζαν τη διανεμημένη ομάδα, οι αξιολογητές της έκβασης και οι αναλυτές των δεδομένων κρατήθηκαν τυφλοποιημένοι για τη διανομή».<sup>156</sup>

«Η τυφλοποίηση και η ισορροπία διατηρήθηκαν αυστηρά, υπογραμμίζοντας στο προσωπικό της παρέμβασης και στους συμμετέχοντες ότι κάθε διατροφή εμμένει στις υγιείς αρχές, και καθεμιά υποστηρίζεται από ορισμένους εμπειρογνώμονες για να είναι ανώτερη όσον αφορά στη μακροπρόθεσμη απώλεια βάρους. Εκτός από τους παρέχοντες τις παρεμβάσεις (διαιτολόγους και συμπεριφορικούς ψυχολόγους), οι ερευνητές και το προσωπικό κρατήθηκαν τυφλοποιημένοι στην ταξινόμηση διατροφής των συμμετεχόντων. Η δοκιμή ενέμεινε στις καθιερωμένες διαδικασίες για να διατηρήσει το διαχωρισμό μεταξύ του προσωπικού που έκανε τις μετρήσεις έκβασης και του προσωπικού που διεκπεραίωσε την παρέμβαση. Τα μέλη του προσωπικού που έλαβαν τις μετρήσεις έκβασης δεν ενημερώθηκαν για την ταξινόμηση της ομάδας διατροφής. Το προσωπικό της παρέμβασης, οι διαιτολόγοι και οι συμπεριφορικοί ψυχολόγοι που διεκπεραίωσαν την παρέμβαση δεν έκαναν τις μετρήσεις έκβασης. Όλοι οι ερευνητές, το προσωπικό και οι συμμετέχοντες κρατήθηκαν καλυμμένοι (masked) στις μετρήσεις της έκβασης και στα αποτελέσματα της δοκιμής».<sup>157</sup>

*Εξήγηση:* Ο όρος «τυφλοποίηση» (blinding) ή «αποκρυπτοποίηση» (masking) αναφέρεται στην παρακράτηση των πληροφοριών για τις ταξινομημένες παρεμβάσεις από τους ανθρώπους που εμπλέκονται στη δοκιμή, οι οποίοι ενδεχομένως να επηρεάζονται από αυτή τη γνώση. Η τυφλοποίηση είναι μια σημαντική προστασία ενάντια στο σφάλμα, ιδιαίτερα κατά την αξιολόγηση υποκειμενικών εκβάσεων.<sup>153</sup>

Ο Benjamin Franklin έχει αναγνωριστεί ως ο πρώτος

που χρησιμοποίησε την τυφλοποίηση σε ένα επιστημονικό πείραμα.<sup>158</sup> Τυφλοποίησε τους συμμετέχοντες, ώστε να μη γνωρίζουν πότε εφαρμόζετο το μερμερισμό (ένα δημοφιλές «θεραπευτικό υγρό» του 18ου αιώνα) και με αυτές τις ενέργειες έδειξε ότι ο μερμερισμός ήταν ψεύτικος. Βασισμένη σε αυτό το πείραμα, η επιστημονική κοινότητα αναγνώρισε τη δύναμη της τυφλοποίησης για να μειώσει το σφάλμα και έχει παραμείνει μια συνήθως χρησιμοποιούμενη στρατηγική στα επιστημονικά πειράματα.

Στο πλαίσιο 4, η ορολογία τυφλοποίησης καθορίζει τις ομάδες ατόμων (δηλαδή συμμετέχοντες, προμηθευτές υπηρεσιών υγείας, συλλέκτες δεδομένων, κριτές εκβάσεων και αναλυτές δεδομένων) που μπορούν πιθανόν να εισάγουν σφάλμα σε μια δοκιμή μέσω της γνώσης των ταξινομήσεων θεραπείας. Οι συμμετέχοντες μπορούν να αποκριθούν διαφορετικά, εάν γνωρίζουν την ταξινόμηση θεραπείας τους (όπως να αποκριθούν ευνοϊκότερα, όταν λαμβάνουν τη νέα θεραπεία).<sup>153</sup> Η έλλειψη της τυφλοποίησης μπορεί επίσης να επηρεάσει τη συμμόρφωση στην παρέμβαση, τη χρήση των συν-παρεμβάσεων και τον κίνδυνο αποχώρησης από τη δοκιμή.

Οι μη τυφλοποιημένοι προμηθευτές υπηρεσιών υγείας μπορεί να εισάγουν παρόμοια σφάλματα, ενώ οι μη τυφλοποιημένοι συλλέκτες δεδομένων μπορεί να αξιολογήσουν τις εκβάσεις διαφορετικά (όπως η συχνότητα ή η χρονική στιγμή), να επαναλάβουν μετρήσεις των μη φυσιολογικών ευρημάτων ή να παρέχουν ενθάρρυνση κατά τη διάρκεια του ελέγχου απόδοσης. Οι μη τυφλοποιημένοι κριτές εκβάσεων μπορεί να αξιολογήσουν διαφορετικά τις υποκειμενικές εκβάσεις, και οι μη τυφλοποιημένοι αναλυτές δεδομένων μπορεί να εισάγουν σφάλμα μέσω της επιλογής αναλυτικών στρατηγικών, όπως η επιλογή των ευνοϊκών χρονικών σημείων ή των εκβάσεων, και με τις αποφάσεις να αποκλείσουν ασθενείς από τις αναλύσεις. Αυτά τα σφάλματα είναι καλά τεκμηριωμένα.<sup>71,153,159–162</sup>

Η τυφλοποίηση, σε αντίθεση με την απόκρυψη διανομής (βλέπε στοιχείο 10), μπορεί να μην είναι πάντα κατάλληλη ή εφικτή. Ένα παράδειγμα είναι μια δοκιμή που συγκρίνει τα επίπεδα πόνου, τα οποία συνδέονται με δειγματοληψία αίματος από το αυτί ή τον αντίχειρα.<sup>163</sup> Η τυφλοποίηση είναι ιδιαίτερα σημαντική όταν τα μέτρα έκβασης προϋποθέτουν κάποια υποκειμενικότητα, όπως η αξιολόγηση του πόνου. Η τυφλοποίηση των συλλεκτών δεδομένων και των κριτών εκβάσεων είναι απίθανο να επηρεάσει αντικειμενικές εκβάσεις, όπως ο θάνατος από οποιαδήποτε αιτία. Ακόμη και τότε, εν τούτοις, η έλλειψη τυφλοποίησης του συμμετέχοντα ή του προμηθευτή υπηρεσιών υγείας μπορεί να οδηγήσει σε άλλα προβλήματα, όπως διαφορετική απώλεια.<sup>164</sup> Σε ορισμένες δοκιμές, ειδικά χειρουργικές δοκιμές, η τυφλοποίηση των συμμετεχόντων

και των χειρουργών συχνά είναι δύσκολη ή αδύνατη, αλλά η τυφλοποίηση των συλλεκτών δεδομένων και των κριτών εκβάσεων είναι συχνά επιτεύξιμη. Για παράδειγμα, οι κακώσεις μπορούν να φωτογραφίζονται πριν και μετά από τη θεραπεία και να αξιολογούνται από έναν εξωτερικό παρατηρητή.<sup>165</sup> Ανεξάρτητα από το εάν η τυφλοποίηση είναι δυνατή, οι συγγραφείς μπορούν και πρέπει πάντα να δηλώσουν ποιος τυφλοποιήθηκε (δηλαδή οι συμμετέχοντες, οι προμηθευτές υπηρεσιών υγείας, οι συλλέκτες δεδομένων και οι κριτές εκβάσεων).

Δυστυχώς, οι συγγραφείς συχνά δεν αναφέρουν εάν χρησιμοποιήθηκε τυφλοποίηση.<sup>166</sup> Για παράδειγμα, στο 51% των 506 δημοσιεύσεων των δοκιμών για την κυστική ίνωση,<sup>167</sup> στο 33% των 196 δοκιμών για τη ρευματοειδή αρθρίτιδα<sup>108</sup> και στο 38% των 68 δοκιμών στη Δερματολογία<sup>8</sup> δεν δήλωσαν εάν χρησιμοποιήθηκε η τυφλοποίηση. Μέχρι οι συγγραφείς των δοκιμών να βελτιώσουν την καταγραφή της τυφλοποίησής τους, οι αναγνώστες θα έχουν δυσκολία στην κρίση της εγκυρότητας των δοκιμών που επιθυμούν να χρησιμοποιήσουν για να καθοδηγηθεί η κλινική πρακτική τους.

Ο όρος «αποκρυπτοποίηση», παρά τυφλοποίηση, χρησιμοποιείται μερικές φορές, για να αποφευχθεί η σύγχυση με την ιατρική κατάσταση έλλειψης της όρασης. Εν τούτοις, η «τυφλοποίηση» υπό τη μεθοδολογική έννοιά της φαίνεται να γίνεται κατανοητή παγκόσμια και είναι αποδεκτή για την καταγραφή κλινικών δοκιμών.<sup>165,168</sup>

### 6.3.15. Στοιχείο 11β. Εάν κρίνεται σχετικό, περιγραφή της ομοιότητας των παρεμβάσεων

*Παράδειγμα:* Η εταιρεία «Εργαστήρια Jamieson» παρείχε 500 mg νιασίνης άμεσης απελευθέρωσης σε ένα λευκό, επίμηκες, διχοτομούμενο δισκίο, σε σχήμα κάψουλας. Επιβεβαιώσαμε ανεξάρτητα το περιεχόμενο του δισκίου, χρησιμοποιώντας υγρή χρωματογραφία υψηλής απόδοσης... Το εικονικό φάρμακο ομοιάζε με το φάρμακο της μελέτης ως προς τη γεύση, το χρώμα και το μέγεθος και περιείχε μικροκρυσταλλική κυτταρίνη, διοξείδιο του πυριτίου, φωσφορικό διασβέστιο, στεατικό μαγνήσιο και στεατικό οξύ.<sup>172</sup>

*Εξήγηση:* Με τον ίδιο τρόπο που αναζητούμε ενδείξεις απόκρυψης για την επιβεβαίωση ότι η ταξινόμηση έγινε με αληθινά τυχαίο τρόπο, αναζητούμε στοιχεία της μεθόδου τυφλοποίησης. Σε δοκιμές με τυφλοποίηση των συμμετεχόντων ή των προμηθευτών υπηρεσιών υγείας, οι συγγραφείς πρέπει να αναφέρουν την ομοιότητα των χαρακτηριστικών των παρεμβάσεων (όπως είναι η εμφάνιση, η γεύση, το άρωμα και ο τρόπος χορήγησης).<sup>35,173</sup>

Κάποιοι συνηγορούν υπέρ του ελέγχου τυφλοποίησης

ρωτώντας στο τέλος της δοκιμής τους συμμετέχοντες και τους προμηθευτές υπηρεσιών υγείας εάν νομίζουν ότι ο συμμετέχων έλαβε την πειραματική παρέμβαση ή την παρέμβαση ελέγχου.<sup>174</sup> Επειδή οι συμμετέχοντες και οι προμηθευτές υπηρεσιών υγείας συνήθως γνωρίζουν αν ο συμμετέχων παρουσίασε την κύρια έκβαση, καθίσταται δύσκολο να καθοριστεί εάν οι απαντήσεις αντανακλούν αποτυχία στη διαδικασία τυφλοποίησης ή ακριβείς υποθέσεις σχετικά με την αποτελεσματικότητα της παρέμβασης.<sup>175</sup> Δεδομένης της αβεβαιότητας που δημιουργούν τέτοιου είδους πληροφορίες αφαιρέσαμε από τη δήλωση CONSORT 2010 την υποστήριξη για αναφορά αυτού του είδους του ελέγχου της τυφλοποίησης. Εν τούτοις, υποστηρίζουμε ότι οι συγγραφείς θα πρέπει να αναφέρουν οποιουδήποτε γνωστού συμβιβασμούς σε σχέση με την τυφλοποίηση. Για παράδειγμα, οι συγγραφείς θα πρέπει να αναφέρουν εάν ήταν αναγκαίο να αρθεί η τυφλοποίηση κάποιου συμμετέχοντα σε οποιοδήποτε σημείο κατά τη διεξαγωγή μιας δοκιμής.

### 6.3.16. Στοιχείο 12α. Στατιστικές μέθοδοι που εφαρμόστηκαν για τη σύγκριση των ομάδων για τις κύριες και τις δευτερεύουσες εκβάσεις

*Παράδειγμα:* «Το κύριο καταληκτικό σημείο ήταν η αλλαγή του βάρους σώματος στη διάρκεια των 20 εβδομάδων της μελέτης στον πληθυσμό με πρόθεση για θεραπεία (intention-to-treat population). Τα δευτερεύοντα καταληκτικά σημεία αποτελεσματικότητας περιελάμβαναν αλλαγή στην περίμετρο μέσης, συστολική και διαστολική αρτηριακή πίεση, επιπολασμό μεταβολικού συνδρόμου... Χρησιμοποιήσαμε ανάλυση συνδιακύμανσης (ANCOVA) για το κύριο καταληκτικό σημείο και για τα δευτερεύοντα καταληκτικά σημεία, όπως περίμετρος μέσης, αρτηριακή πίεση και τα αναφερόμενα από τον ασθενή αποτελέσματα έκβασης' αυτή συμπληρώθηκε από μια ανάλυση επαναλαμβανόμενων μετρήσεων. Το μοντέλο της ANCOVA περιελάμβανε τη θεραπεία, τη χώρα και το φύλο ως σταθερές επιδράσεις, καθώς και το σωματικό βάρος κατά την τυχαίωση ως συμμεταβλητή. Είχαμε ως στόχο να αξιολογήσουμε κατά πόσον τα δεδομένα παρείχαν ενδείξεις ανωτερότητας κάθε δόσης λιραγλουτίδης σε σχέση με το εικονικό φάρμακο (πρωταρχικός στόχος) και σε σχέση με την ορλιστάτη (δευτερεύων στόχος)».<sup>176</sup>

*Εξήγηση:* Τα δεδομένα μπορούν να αναλυθούν με διάφορους τρόπους, μερικοί από τους οποίους ενδέχεται να μην είναι αυστηρά κατάλληλοι σε μια συγκεκριμένη περίπτωση. Είναι απαραίτητο να διευκρινιστεί ποια στατιστική δοκιμασία χρησιμοποιήθηκε για κάθε ανάλυση και πιθανόν να απαιτείται περαιτέρω διευκρίνιση στο τμήμα των αποτελεσμάτων της δημοσίευσης. Η αρχή που θα



πρέπει να ακολουθηθεί είναι: «Περιγράψτε τις στατιστικές μεθόδους με τόση λεπτομέρεια, ώστε ένας πεπειραμένος αναγνώστης με πρόσβαση στα αρχικά δεδομένα να μπορεί να επαληθεύσει τα αναφερόμενα αποτελέσματα» ([www.icmje.org](http://www.icmje.org)). Επίσης, είναι σημαντικό να περιγραφούν λεπτομέρειες της στατιστικής ανάλυσης, όπως είναι η ανάλυση πρόθεσης για θεραπεία (βλέπε πλαίσιο 6).

Σχεδόν όλες οι μέθοδοι ανάλυσης παρέχουν μια εκτίμηση του θεραπευτικού αποτελέσματος, η οποία αποτελεί μια αντίθεση μεταξύ των αποτελεσμάτων στις ομάδες σύγκρισης. Οι συγγραφείς θα πρέπει να τη συνοδεύουν με ένα διάστημα εμπιστοσύνης για την εκτιμώμενη επίδραση, το οποίο υποδεικνύει ένα κεντρικό εύρος της αβεβαιότητας για το πραγματικό θεραπευτικό αποτέλεσμα. Το διάστημα εμπιστοσύνης μπορεί να ερμηνευτεί ως το εύρος τιμών του θεραπευτικού αποτελέσματος, το οποίο είναι συμβατό με τα παρατηρούμενα δεδομένα. Είναι σύνηθες να παρουσιάζεται ένα διάστημα εμπιστοσύνης 95%, το οποίο παρέχει το εύρος που αναμένεται να περιλαμβάνει την πραγματική τιμή στις 95 από τις 100 παρόμοιες μελέτες.

Τα ευρήματα των μελετών μπορούν, ακόμη, να αξιολογηθούν σχετικά με τη στατιστική τους σημαντικότητα. Η τιμή  $p$  αντιπροσωπεύει την πιθανότητα, τα παρατηρούμενα δεδομένα (ή ένα πιο ακραίο αποτέλεσμα) να έχουν προκύψει κατά τύχη, όταν οι παρεμβάσεις δεν διαφέρουν αληθινά μεταξύ τους. Οι πραγματικές τιμές του  $p$  (για παράδειγμα,  $p=0,003$ ) είναι έντονα προτιμότερες από ασαφείς αναφορές ορίων, όπως  $p<0,05$ .<sup>48,177</sup>

Οι τυποποιημένες μέθοδοι ανάλυσης υποθέτουν ότι τα δεδομένα είναι «ανεξάρτητα». Για τις ελεγχόμενες δοκιμές αυτό συνήθως σημαίνει ότι υπάρχει μία παρατήρηση για κάθε συμμετέχοντα. Η διαχείριση πολλαπλών παρατηρήσεων από ένα συμμετέχοντα ως ανεξάρτητα δεδομένα αποτελεί σοβαρό σφάλμα. Τέτοια δεδομένα παράγονται όταν τα αποτελέσματα μπορούν να μετρηθούν σε διαφορετικά σημεία του σώματος, όπως γίνεται στην Οδοντιατρική και στη Ρευματολογία. Η ανάλυση των δεδομένων θα πρέπει να στηρίζεται στην εφ' άπαξ καταμέτρηση κάθε συμμετέχοντα<sup>178,179</sup> ή θα πρέπει να γίνεται χρησιμοποιώντας πιο πολύπλοκες στατιστικές διαδικασίες.<sup>180</sup> Εσφαλμένη ανάλυση πολλαπλών παρατηρήσεων ανά συμμετέχοντα παρατηρήθηκε σε 123 (63%) από 196 δοκιμές στη ρευματοειδή αρθρίτιδα.<sup>108</sup>

### 6.3.17. Στοιχείο 12β. Μέθοδοι για πρόσθετες αναλύσεις, όπως αναλύσεις υποομάδων και σταθμισμένες αναλύσεις

*Παραδείγματα:* «Οι αναλογίες των ασθενών που ανταποκρίνονταν, συγκρίθηκαν μεταξύ των ομάδων θεραπείας με τη δοκιμασία Mantel Haenszel  $\chi^2$ , σταθμισμένη για τη μεταβλητή διαστρωμάτωσης, τη χρήση μεθοτρεξάτης».<sup>103</sup>

Η προκαθορισμένη ανάλυση υποομάδων πραγματοποιήθηκε με τη χρήση διαστρωματοποιημένων μοντέλων αναλογικού κινδύνου κατά Cox, σύμφωνα με την ταξινόμηση (ταξινομήσεις) στην αντιοξειδωτική θεραπεία, την παρουσία ή την απουσία προηγούμενης καρδιαγγειακής νόσου, τη διατροφική πρόσληψη φυλλικού οξέος, το κάπνισμα, το διαβήτη, την ασπιρίνη, την ορμονική θεραπεία και τη χρήση πολυβιταμινών. Αυτές οι αναλύσεις χρησιμοποιήσαν την εκτίμηση των αρχικών βασικών τιμών κατά την έκθεση και περιορίστηκαν σε συμμετέχοντες για τους οποίους δεν υπήρχαν ελλιπή στοιχεία στις αρχικές τιμές, στα δεδομένα των υποομάδων».<sup>181</sup>

*Εξήγηση:* Όπως και στην περίπτωση των κύριων αναλύσεων, η μέθοδος της ανάλυσης υποομάδων πρέπει να είναι σαφώς οριζόμενη. Οι πιο ισχυρές αναλύσεις είναι αυτές που αναζητούν διαφορά στο θεραπευτικό αποτέλεσμα σε συμπληρωματικές υποομάδες (για παράδειγμα, νεότεροι και πρεσβύτεροι συμμετέχοντες), μια σύγκριση γνωστή και ως έλεγχος αλληλεπίδρασης.<sup>182,183</sup> Μια συνήθης αλλά αποπροσανατολιστική προσέγγιση είναι η σύγκριση των τιμών του  $p$  για ξεχωριστές αναλύσεις του θεραπευτικού αποτελέσματος σε κάθε ομάδα. Δεν είναι σωστό να συμπεραίνουμε επίδραση υποομάδας (αλληλεπίδραση) από μια σημαντική και μια μη σημαντική τιμή του  $p$ .<sup>184</sup> Τέτοια συμπερασματολογία έχει αυξημένο ποσοστό ψευδώς θετικών αποτελεσμάτων.

Εξ αιτίας του υψηλού κινδύνου για νόθα ευρήματα, οι αναλύσεις των υποομάδων συχνά δεν ενθαρρύνονται.<sup>14,185</sup> Συγκρίσεις υποομάδων εκ των υστέρων –αναλύσεις που διενεργούνται μετά από την εξέταση των δεδομένων– είναι ιδιαίτερα πιθανό να μην επιβεβαιωθούν από περαιτέρω μελέτες. Τέτοιες αναλύσεις δεν έχουν μεγάλη αξιοπιστία.

Σε κάποιες μελέτες γίνεται στάθμιση για ανισορροπίες στα χαρακτηριστικά των συμμετεχόντων, χρησιμοποιώντας έναν τύπο ανάλυσης πολλαπλής παλινδρόμησης. Αν και η ανάγκη για στάθμιση είναι πολύ μικρότερη στις ΤΚΔ από τις επιδημιολογικές μελέτες, μια σταθμισμένη ανάλυση μπορεί να είναι λογική, ειδικά εάν μία ή περισσότερες μεταβλητές θεωρούνται προγνωστικές.<sup>186</sup> Ιδανικά, οι σταθμισμένες αναλύσεις πρέπει να καθορίζονται στο πρωτόκολλο της μελέτης (βλέπε στοιχείο 24). Για παράδειγμα, η στάθμιση συστήνεται συχνά για κάθε μεταβλητή διαστρωμάτωσης (βλέπε στοιχείο 8β) επί της αρχής ότι η στρατηγική της ανάλυσης πρέπει να ακολουθεί το σχεδιασμό. Στις ΤΚΔ, η απόφαση για προσαρμογή δεν πρέπει να καθορίζεται από το εάν οι διαφορές στις αρχικές τιμές ήταν στατιστικά σημαντικές (βλέπε στοιχείο 16).<sup>183,187</sup> Η λογική για κάθε σταθμισμένη ανάλυση και οι στατιστικές μέθοδοι που χρησιμοποιήθηκαν πρέπει να καθορίζονται.



Οι συγγραφείς πρέπει να αποσαφηνίζουν την επιλογή των μεταβλητών βάσει των οποίων έγινε η στάθμιση, να υποδεικνύουν πώς χειρίστηκαν τις συνεχείς μεταβλητές και να καθορίζουν αν η ανάλυση σχεδιάστηκε ή προέκυψε από τα δεδομένα.<sup>188</sup> Ανασκοπήσεις δημοσιευμένων μελετών δείχνουν ότι η αναφορά σταθμισμένων αναλύσεων είναι ανεπαρκής σε σχέση με όλα αυτά τα ζητήματα.<sup>188-191</sup>

**7. ΑΠΟΤΕΛΕΣΜΑΤΑ**

**7.1. Στοιχείο 13. Ροή συμμετεχόντων (ένα διάγραμμα συστήνεται έντονα)**

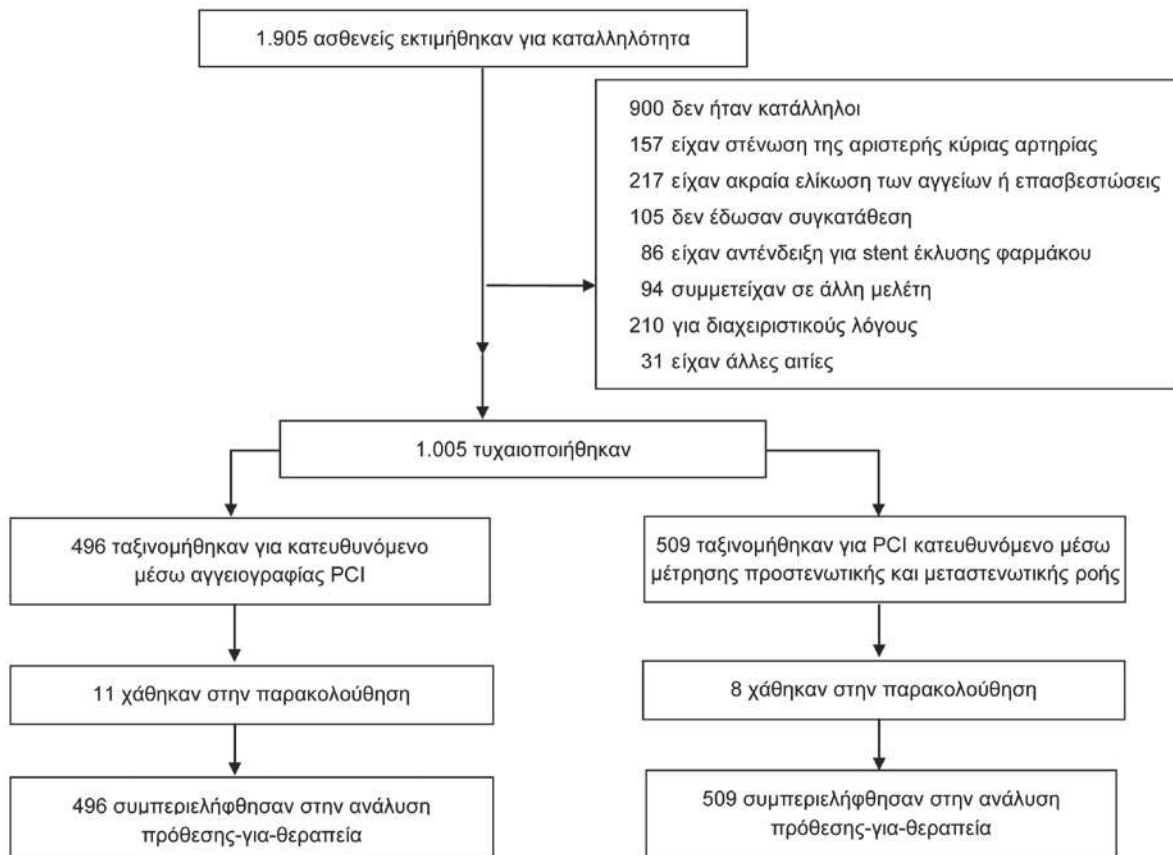
**7.1.1. Στοιχείο 13α. Για κάθε ομάδα, ο αριθμός των συμμετεχόντων, που ταξινομήθηκαν τυχαία, οι οποίοι έλαβαν την προορισμένη παρέμβαση και αναλύθηκαν για την κύρια έκβαση**

*Παραδείγματα:* Βλέπε εικόνες 2 και 3.

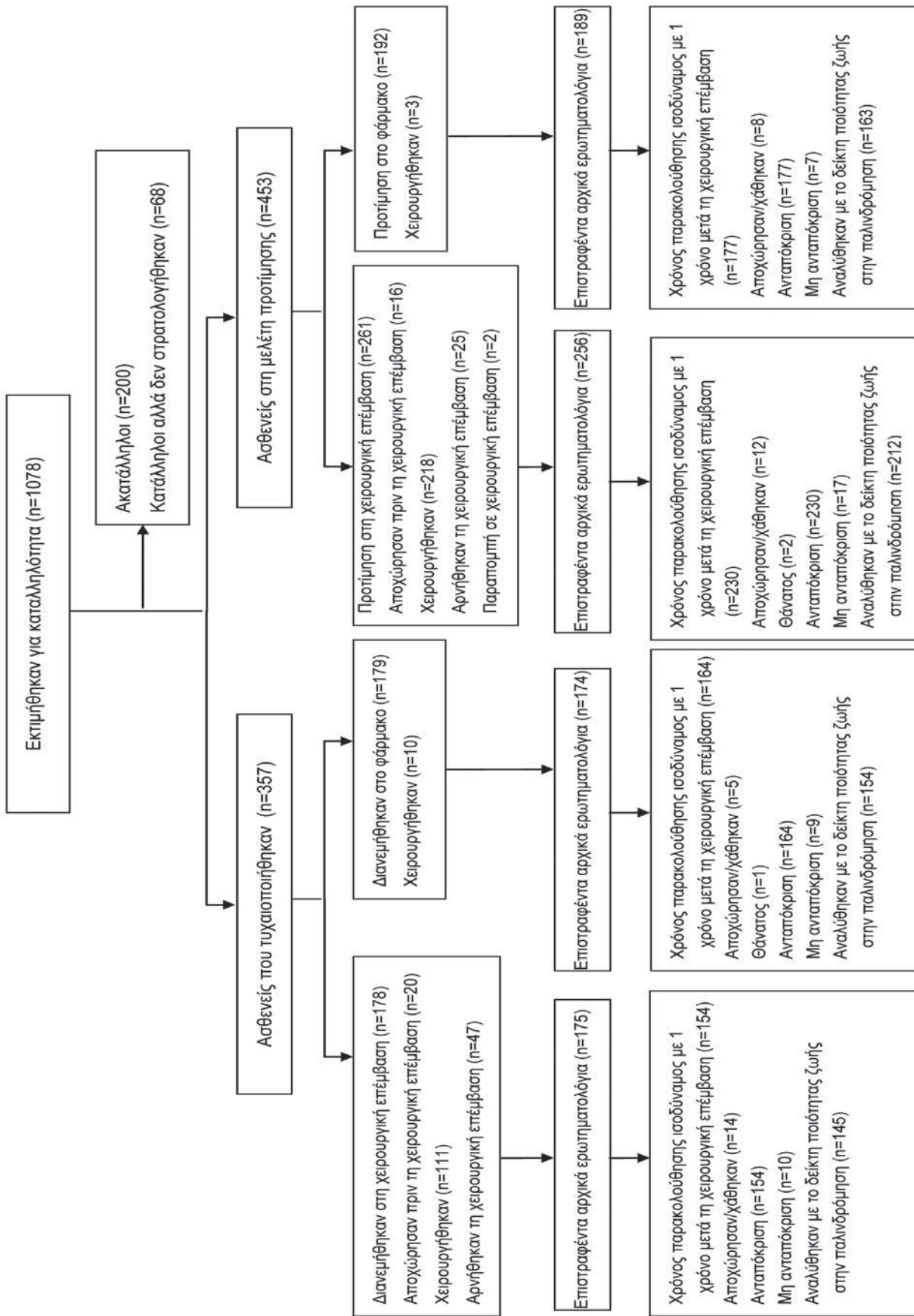
*Εξήγηση:* Ο σχεδιασμός και η διενέργεια κάποιων ΤΚΔ είναι σαφείς, και η ροή των συμμετεχόντων, ιδιαίτερα όπου δεν υπάρχουν απώλειες κατά την παρακολούθηση

ή αποκλεισμοί, σε κάθε φάση της μελέτης, μπορεί να περιγραφεί ικανοποιητικά σε μερικές προτάσεις. Σε μερικές περίπλοκες μελέτες, μπορεί να είναι δύσκολο για τους αναγνώστες να διακρίνουν αν και γιατί κάποιοι συμμετέχοντες δεν έλαβαν τη θεραπεία, όπως διανεμήθηκε, αν χάθηκαν κατά την παρακολούθηση, ή αν αποκλείστηκαν από την ανάλυση.<sup>51</sup> Οι πληροφορίες αυτές είναι κρίσιμες για διάφορους λόγους. Οι συμμετέχοντες που αποκλείστηκαν μετά από τη διανομή είναι απίθανο να είναι αντιπροσωπευτικοί του συνόλου των συμμετεχόντων στη μελέτη. Για παράδειγμα, οι ασθενείς μπορεί να μην είναι διαθέσιμοι για εκτίμηση κατά την παρακολούθηση, γιατί εμφάνισαν μια οξεία παρόξυνση της νόσου ή βλαβερές συνέπειες της θεραπείας.<sup>22,192</sup>

Η απώλεια, ως αποτέλεσμα της μη παρακολούθησης, στοιχείο που συχνά δεν είναι δυνατόν να αποφευχθεί, πρέπει να διακριθεί από τον αποκλεισμό εκ μέρους του ερευνητή για λόγους όπως η ακαταλληλότητα, η αποχώρηση από τη θεραπεία και η κακή προσκόλληση στο πρωτόκολλο της δοκιμής. Λανθασμένα συμπεράσματα μπορούν να εξαχθούν αν οι συμμετέχοντες αποκλείονται από τις αναλύσεις, και οι ανισορροπίες ανάμεσα στις ομάδες, λόγω τέτοιων απο-



**Εικόνα 2.** Διάγραμμα ροής μιας πολυκεντρικής δοκιμής της μέτρησης προστενωτικής και μεταστενωτικής στεφανιαίας αιματικής ροής έναντι αγγειογραφίας και την καθοδήγηση διαδερμικής στεφανιαίας παρέμβασης (percutaneous coronary intervention, PCI) (προσαρμοσμένο από τους Torino et al<sup>213</sup>). Το διάγραμμα περιλαμβάνει λεπτομερείς πληροφορίες για τους ασθενείς που αποκλείστηκαν.



**Εικόνα 3.** Διάγραμμα ροής της χειρουργικής ελάχιστης παρεμβατικότητας σε σχέση με τη φαρμακευτική αντιμετώπιση της χρόνιας γαστρο-οισοφαγικής παλινδρομικής νόσου (προσαρμοσμένο από τους Grant et al<sup>196</sup>). Το διάγραμμα δείχνει μια παράλληλη μη τυχαιοποιημένη ομάδα προτίμησης.

κλεισμών, είναι ιδιαίτερα ενδεικτικές σφαλμάτων.<sup>192-194</sup> Γι' αυτόν το λόγο, πληροφορίες σχετικές με το αν οι ερευνητές συμπεριέλαβαν στην ανάλυση όλους τους συμμετέχοντες που τυχαιοποιήθηκαν και εάν αυτοί αναλύθηκαν στις ομάδες στις οποίες είχαν αρχικά διανεμηθεί (ανάλυση πρόθεσης για θεραπεία, βλέπε στοιχείο 16 και πλαίσιο 6) είναι ιδιαίτερης σημασίας. Γνωρίζοντας τον αριθμό των συμμετεχόντων που δεν έλαβαν την παρέμβαση, όπως διανεμήθηκαν ή δεν ολοκλήρωσαν τη θεραπεία, επιτρέπεται στον αναγνώστη να εκτιμήσει το βαθμό πιθανής υποεκτίμησης της εκτιμώμενης αποτελεσματικότητας της θεραπείας σε σχέση με τις ιδανικές συνθήκες.

Αν είναι διαθέσιμος, θα πρέπει να αναφέρεται και ο αριθμός των ατόμων οι οποίοι εκτιμήθηκαν για καταλληλότητα. Αν και ο αριθμός αυτός έχει σχέση μόνο με την εξωτερική εγκυρότητα και είναι αναμφισβήτητο λιγότερο σημαντικός από άλλες μετρήσεις,<sup>195</sup> είναι ένας χρήσιμος δείκτης του αν οι συμμετέχοντες στη μελέτη είναι πιθανό να είναι αντιπροσωπευτικοί του συνόλου των κατάλληλων συμμετεχόντων.

Μια ανασκόπηση ΤΚΔ που δημοσιεύτηκε το 1998 σε

πέντε κορυφαία γενικά ιατρικά περιοδικά και ιατρικά περιοδικά Εσωτερικής Παθολογίας εντόπισε ότι η αναφορά της ροής των συμμετεχόντων ήταν συχνά ατελής, ειδικά σε σχέση με τον αριθμό των συμμετεχόντων που έλαβαν την παρέμβαση, όπως είχαν διανεμηθεί, και τον αριθμό των συμμετεχόντων οι οποίοι απωλέστηκαν κατά την παρακολούθηση.<sup>51</sup> Ακόμη και πληροφορίες τόσο βασικές όσο ο αριθμός των συμμετεχόντων που υποβλήθηκαν σε τυχαιοποίηση και ο αριθμός αυτών που αποκλείστηκαν από την ανάλυση δεν ήταν διαθέσιμες στο 20% των άρθρων.<sup>51</sup> Η αναφορά ήταν σημαντικά λεπτομερέστερη σε άρθρα που περιελάμβαναν διάγραμμα ροής των συμμετεχόντων στη μελέτη, όπως προτείνεται από το CONSORT. Η μελέτη αυτή ενημερώνει για το σχεδιασμό του αναθεωρημένου διαγράμματος ροής στην αναθεωρημένη δήλωση CONSORT.<sup>52-54</sup> Το προτεινόμενο πρότυπο εμφανίζεται στην εικόνα 1 και οι απαιτούμενες μετρήσεις περιγράφονται με λεπτομέρεια στον πίνακα 3.

Μερικές πληροφορίες, όπως ο αριθμός των ατόμων που αξιολογούνται για την καταλληλότητα, μπορεί να μην είναι πάντα γνωστές<sup>14</sup> και, ανάλογα με τη φύση της

**Πίνακας 3.** Πληροφορίες απαραίτητες για την τεκμηρίωση της ροής των συμμετεχόντων σε κάθε στάδιο μιας τυχαιοποιημένης δοκιμής.

Στάδιο	Αριθμός ατόμων που συμπεριελήφθησαν	Αριθμός ατόμων που δεν συμπεριελήφθησαν ή αποκλείστηκαν	Σκεπτικό
Εγγραφή	Άτομα που αξιολογήθηκαν για πιθανή εγγραφή	Άτομα που δεν πληρούσαν τα κριτήρια εισόδου ή πληρούσαν τα κριτήρια εισόδου αλλά αρνήθηκαν την εγγραφή	Οι αριθμοί αυτοί υποδεικνύουν αν οι συμμετέχοντες στη δοκιμή ήταν πιθανόν αντιπροσωπευτικοί όλων των ασθενών που αξιολογήθηκαν· είναι σχετικοί με την εκτίμηση της εξωτερικής εγκυρότητας και συχνά δεν είναι διαθέσιμοι
Τυχαιοποίηση	Συμμετέχοντες, τυχαία ταξινομημένοι		Καίριος αριθμός για καθορισμό του μεγέθους της δοκιμής και για την εκτίμηση αν μια δοκιμή αναλύθηκε σύμφωνα με την πρόθεση-για-θεραπεία
Διανομή θεραπείας	Συμμετέχοντες οι οποίοι ολοκλήρωσαν τη θεραπεία σύμφωνα με τη διανομή, κατά ομάδα της μελέτης	Συμμετέχοντες οι οποίοι δεν ολοκλήρωσαν τη θεραπεία σύμφωνα με τη διανομή, κατά ομάδα της μελέτης	Σημαντικοί αριθμοί για την εκτίμηση της εσωτερικής εγκυρότητας και της ερμηνείας των αποτελεσμάτων. Πρέπει να δίνονται οι αιτίες για μη λήψη της θεραπείας σύμφωνα με τη διανομή
Παρακολούθηση	Συμμετέχοντες οι οποίοι ολοκλήρωσαν τη θεραπεία σύμφωνα με τη διανομή, κατά ομάδα της μελέτης Συμμετέχοντες οι οποίοι ολοκλήρωσαν την παρακολούθηση, όπως σχεδιάστηκε, κατά ομάδα της μελέτης	Συμμετέχοντες οι οποίοι δεν ολοκλήρωσαν τη θεραπεία σύμφωνα με τη διανομή, κατά ομάδα της μελέτης Συμμετέχοντες οι οποίοι δεν ολοκλήρωσαν την παρακολούθηση, όπως σχεδιάστηκε, κατά ομάδα της μελέτης	Σημαντικοί αριθμοί για την εκτίμηση της εσωτερικής εγκυρότητας και της ερμηνείας των αποτελεσμάτων. Πρέπει να παρατίθενται οι αιτίες για μη ολοκλήρωση της θεραπείας ή της παρακολούθησης σύμφωνα με τη διανομή
Ανάλυση	Συμμετέχοντες οι οποίοι συμπεριελήφθησαν στην κύρια ανάλυση, κατά ομάδα της μελέτης	Συμμετέχοντες οι οποίοι αποκλείστηκαν από την κύρια ανάλυση, κατά ομάδα της μελέτης	Σημαντικός αριθμός για την αξιολόγηση αν μια δοκιμή αναλύθηκε σύμφωνα με την πρόθεση-για-θεραπεία. Πρέπει να αναφέρονται οι αιτίες για αποκλεισμό των συμμετεχόντων

μελέτης, μερικές μετρήσεις μπορεί να είναι πιο σχετικές από άλλες. Μερικές φορές μπορεί να είναι χρήσιμο ή αναγκαίο να προσαρμόζεται η δομή του διαγράμματος ροής σε μια συγκεκριμένη δοκιμή. Σε κάποιες περιπτώσεις, ενδέχεται να προστεθούν και άλλες χρήσιμες πληροφορίες. Για παράδειγμα, το διάγραμμα ροής μιας δοκιμής παράλληλων ομάδων χειρουργικής ελάχιστης παρεμβατικότητας συγκρινόμενη με την ιατρική θεραπεία για τη χρόνια γαστροοισοφαγική παλινδρόμηση μπορεί να περιλαμβάνει, επίσης, μια παράλληλη, μη τυχαιοποιημένη ομάδα προτίμησης (εικ. 3).<sup>196</sup>

Η ακριβής ροή και το περιεχόμενο του διαγράμματος ροής μπορεί να διαφέρει ανάλογα με τα ιδιαίτερα χαρακτηριστικά της μελέτης. Για παράδειγμα, πολλές δοκιμές χειρουργικής ή εμβολιασμών δεν περιλαμβάνουν το ενδεχόμενο της διακοπής. Παρ' όλο που το CONSORT συστήνει έντονα τη χρήση αυτού του γραφικού εργαλείου για τη μετάδοση της ροής των συμμετεχόντων κατά τη διάρκεια ολόκληρης της μελέτης, δεν υπάρχει ειδική, προκαθορισμένη μορφή.

### 7.1.2. Στοιχείο 13β. Για κάθε ομάδα, απώλειες και αποκλεισμοί μετά από την τυχαιοποίηση, σε συνδυασμό με αιτίες

*Παράδειγμα:* «Υπήρχε μόνο μία παράκαμψη από το πρωτόκολλο που αφορούσε σε μία γυναίκα στην ομάδα μελέτης. Είχε μια μη φυσιολογική πυελική μέτρηση και προγραμματίστηκε για επιλεκτική καισαρική τομή. Εν τούτοις, ο περιθάλπων μαιευτήρας, που παρευρισκόταν, έκρινε τη δοκιμή του φυσιολογικού τοκετού αποδεκτή. Η καισαρική τομή ακολούθησε, όταν δεν υπήρξε πρόοδος στο πρώτο στάδιο του τοκετού».<sup>197</sup>

«Η καταγραφή οδήγησε σε ανάκληση 9 κέντρων, στα οποία η ύπαρξη μερικών ασθενών δεν μπορούσε να επιβεβαιωθεί, ή είχαν συμβεί και άλλες σοβαρές παραβάσεις της καλής κλινικής πρακτικής».<sup>198</sup>

*Εξήγηση:* Μερικές παρακάμψεις από το πρωτόκολλο μπορεί να αναφέρονται στο διάγραμμα ροής (βλέπε στοιχείο 13α)· για παράδειγμα, οι συμμετέχοντες οι οποίοι δεν έλαβαν την προτιθέμενη παρέμβαση. Αν οι συμμετέχοντες αποκλείονταν μετά από την τυχαιοποίηση (σε αντίθεση με την αρχή της πρόθεσης για θεραπεία), επειδή βρέθηκαν να μην πληρούν τα κριτήρια καταλληλότητας (βλέπε στοιχείο 16), πρέπει να περιλαμβάνονται στο διάγραμμα ροής. Η χρήση του όρου «παράκαμψη από το πρωτόκολλο» σε δημοσιευμένα άρθρα δεν είναι επαρκής για να δικαιολογήσει τον αποκλεισμό των συμμετεχόντων μετά από την τυχαιοποίηση. Η φύση της παράκαμψης του

πρωτοκόλλου και ο ακριβής λόγος για τον αποκλεισμό των συμμετεχόντων μετά από την τυχαιοποίηση θα πρέπει να αναφέρονται πάντα.

### 7.2. Στοιχείο 14α. Ημερομηνίες που καθορίζουν τις περιόδους στρατολόγησης και παρακολούθησης

*Παράδειγμα:* «Ηλικιακά κατάλληλοι συμμετέχοντες στρατολογήθηκαν... από το Φεβρουάριο του 1993 έως το Σεπτέμβριο του 1994... Οι συμμετέχοντες παρίσταντο στις επισκέψεις της κλινικής τη στιγμή της τυχαιοποίησης (baseline, αρχική μέτρηση) και κάθε 6 μήνες για 3 χρόνια».<sup>199</sup>

*Εξήγηση:* Το να είναι γνωστό πότε έλαβε χώρα η μελέτη και ποια περίοδο στρατολογήθηκαν οι συμμετέχοντες, θέτει τη μελέτη σε ιστορικό πλαίσιο. Οι ιατρικές και οι χειρουργικές θεραπείες, περιλαμβανομένων των ταυτόχρονων θεραπειών, εξελίσσονται συνεχώς και μπορεί να επηρεάσουν τη συνήθη φροντίδα που παρέχεται στους συμμετέχοντες κατά τη διάρκεια μιας δοκιμής. Γνωρίζοντας το ρυθμό με τον οποίο στρατολογήθηκαν οι συμμετέχοντες μπορεί να είναι ακόμη χρήσιμο, ιδιαίτερα σε άλλους ερευνητές.

Η διάρκεια της παρακολούθησης δεν είναι πάντα προκαθορισμένη μετά από την τυχαιοποίηση. Σε πολλές ΤΚΔ, στις οποίες η έκβαση είναι ο χρόνος προς ένα συμβάν, η παρακολούθηση όλων των συμμετεχόντων ολοκληρώνεται σε μια συγκεκριμένη ημερομηνία. Η ημερομηνία αυτή θα πρέπει να παρέχεται, ενώ είναι ακόμη χρήσιμο να αναφέρεται η ελάχιστη, η μέγιστη και η διάμεση διάρκεια παρακολούθησης.<sup>200,201</sup>

Σε μια ανασκόπηση δημοσιεύσεων σε ογκολογικά περιοδικά, οι οποίες χρησιμοποίησαν ανάλυση επιβίωσης και οι περισσότερες δεν ήταν ΤΚΔ,<sup>201</sup> βρέθηκε ότι σχεδόν στο 80% (104 από τις 132 δημοσιεύσεις) συμπεριελήφθησαν η αρχική και η τελική ημερομηνία εγγραφής των ασθενών, αλλά μόνο σε ποσοστό 24% (32 από τις 132 αναφορές) αναφέρθηκε επίσης η ημερομηνία κατά την οποία σταμάτησε η παρακολούθηση.

### 7.3. Στοιχείο 14β. Γιατί η δοκιμή τερματίστηκε ή διακόπηκε;

*Παράδειγμα:* «Κατά τη στιγμή της ενδιάμεσης ανάλυσης, η συνολική παρακολούθηση περιελάμβανε 63% του υπολογιζόμενου αριθμού των ασθενο-ετών (patient-years), τα οποία θα είχαν συλλεχθεί στο τέλος της μελέτης, οδηγώντας σε μια οριακή τιμή 0,0095, όπως καθορίζεται από τη μέθοδο διόρθωσης του επιπέδου σημαντικότητας σύμφωνα με το κριτήριο Lan-De Mets... Στην ενδιάμεση



ανάλυση, ο RR (relative risk, σχετικός κίνδυνος) ήταν 0,37 στην ομάδα παρέμβασης, σε σχέση με την ομάδα ελέγχου, με μια τιμή p που αντιστοιχούσε σε 0,00073, κάτω από την οριακή τιμή. Η επιτροπή ασφάλειας και επιτήρησης δεδομένων συμβούλευσε τους ερευνητές να διακόψουν τη δοκιμή και να παρέχουν περιτομή στους ασθενείς της ομάδας ελέγχου, από τους οποίους στη συνέχεια ζητήθηκε να προσέλθουν στο ερευνητικό κέντρο, όπου συστήθηκε και προτάθηκε MC (medical circumcision, ιατρική περιτομή)... Επειδή η μελέτη διακόπηκε, κάποιιοι από τους συμμετέχοντες δεν είχαν ολοκληρωμένη παρακολούθηση σε εκείνη την ημερομηνία, και οι επισκέψεις τους οι οποίες δεν είχαν ολοκληρωθεί, περιγράφηκαν ως «σχεδιαζόμενες» στο άρθρο».<sup>202</sup>

«Τον Ιανουάριο του 2000, προβλήματα στην προμήθεια των εμβολίων κατέστησαν αναγκαία την προσωρινή αντικατάσταση, σε εθνικό επίπεδο, του ολοκληρωτικού συστατικού του συζευγμένου εμβολίου DPT/Hib με το ακυτταρικό αντικοκκυτικό εμβόλιο. Επειδή το εμβόλιο αυτό έχει ένα διαφορετικό τοπικό προφίλ αντιδραστικότητας, αποφασίσαμε να διακόψουμε πρώιμα τη δοκιμή».<sup>203</sup>

*Εξήγηση:* Αναμφισβήτητα, οι διεξάγοντες τη δοκιμή, οι οποίοι αυθαίρετα διενεργούν μη προγραμματισμένες ενδιάμεσες αναλύσεις, μετά από την εγγραφή πολύ λίγων συμβάντων, χωρίς τη χρήση στατιστικών οδηγιών, διατρέχουν έναν υψηλό κίνδυνο να «πιάσουν» τα δεδομένα σε ένα τυχαίο άκρο, το οποίο πιθανόν αντιπροσωπεύει μια μεγάλη υπερεκτίμηση του οφέλους της θεραπείας.<sup>204</sup>

Οι αναγνώστες ενδεχομένως να εξάγουν ασθενέστερα συμπεράσματα από μια δοκιμή η οποία διακόπηκε με τρόπο που βασίστηκε στα δεδομένα, σε σύγκριση με μια άλλη, η οποία αναφέρει τα ευρήματα μετά την επίτευξη ενός σκοπού, ανεξάρτητα από τα αποτελέσματα. Γι' αυτό, οι ΤΚΔ θα πρέπει να επισημαίνουν γιατί η δοκιμή τελείωσε (βλέπε πλαίσιο 5). Η δημοσίευση θα πρέπει ακόμη να αποκαλύπτει παράγοντες εξωγενείς της δοκιμής, οι οποίοι επηρέασαν την απόφαση για τον τερματισμό της δοκιμής, και ποιος αποφάσισε τη διακοπή της δοκιμής, περιλαμβάνοντας την αναφορά του ρόλου τον οποίο διαδραμάτισε ο φορέας χρηματοδότησης στις διαβουλεύσεις και στην απόφαση για τη διακοπή της δοκιμής.<sup>134</sup>

Μια συστηματική ανασκόπηση 143 ΤΚΔ, οι οποίες διακόπηκαν νωρίτερα του προγραμματισμένου λόγω οφέλους, βρήκε ότι οι δοκιμές αυτές ανέφεραν πως διακόπηκαν μετά από την εγγραφή 66 συμβάντων κατά διάμεσο όρο και εκτιμήθηκε ένας διάμεσος σχετικός κίνδυνος της τάξης του 0,47, καθώς και μια ισχυρή σχέση μεταξύ του αριθμού των συμβάντων, που εγγράφηκαν, και του μεγέθους της επίδρασης, με τις μικρότερες δοκιμές με τα λιγότερα συμβάντα να

αποδίδουν μεγαλύτερα θεραπευτικά αποτελέσματα (λόγος αναλογιών 31, 95% διαστήματα εμπιστοσύνης 12–82).<sup>134</sup> Ενώ ένας αυξανόμενος αριθμός δοκιμών δημοσιευμένων σε ιατρικά περιοδικά με υψηλό συντελεστή απήχησης αναφέρουν ότι διέκοψαν νωρίς, μόνο 0,1% των δοκιμών αναφέρουν ότι σταμάτησαν νωρίς λόγω οφέλους, γεγονός το οποίο έρχεται σε αντίθεση με εκτιμήσεις από μελέτες προσομοίωσης<sup>205</sup> και έρευνες από επιτροπές ασφάλειας και επιτήρησης δεδομένων.<sup>206</sup> Συνεπώς, πολλές μελέτες, οι οποίες εγγράφουν λίγους συμμετέχοντες και αναφέρουν μεγάλα θεραπευτικά αποτελέσματα, ενδέχεται να έχουν σταματήσει νωρίτερα από τον αρχικό σχεδιασμό και παρέρειψαν να αναφέρουν αυτή την πράξη.

#### 7.4. Στοιχείο 15. Ένας πίνακας στον οποίο να φαίνονται τα βασικά δημογραφικά και κλινικά χαρακτηριστικά για κάθε ομάδα

*Παράδειγμα:* Βλέπε πίνακα 4.

*Εξήγηση:* Παρ' όλο που τα κριτήρια καταλληλότητας (βλέπε στοιχείο 4α) υποδεικνύουν ποιος ήταν κατάλληλος για τη δοκιμή, είναι ακόμη σημαντικό να είναι γνωστά τα χαρακτηριστικά των συμμετεχόντων, οι οποίοι πραγματικά περιελήφθησαν. Οι πληροφορίες αυτές επιτρέπουν στους αναγνώστες, ειδικά στους κλινικούς, να κρίνουν πόσο σχετικά μπορεί να είναι τα αποτελέσματα της δοκιμής σε ένα συγκεκριμένο ασθενή.

Οι κλινικές δοκιμές στοχεύουν στη σύγκριση ομάδων συμμετεχόντων, οι οποίες διαφέρουν μόνο σε σχέση με την παρέμβαση (θεραπεία). Αν και η κατάλληλη τυχαία ταξινόμηση προφυλάσσει από σφάλματα επιλογής, δεν παρέχει εγγυήσεις ότι οι ομάδες θα είναι ισοδύναμες στα αρχικά τους χαρακτηριστικά. Κάθε διαφορά στα αρχικά βασικά χαρακτηριστικά είναι, εν τούτοις, το αποτέλεσμα της τύχης παρά των σφαλμάτων.<sup>32</sup> Οι ομάδες της μελέτης θα πρέπει να συγκρίνονται ως προς τα αρχικά σημαντικά δημογραφικά και κλινικά χαρακτηριστικά, έτσι ώστε οι αναγνώστες να εκτιμούν πόσο παρόμοιες ήταν. Τα αρχικά δεδομένα είναι ιδιαίτερα πολύτιμα για εκβάσεις που μπορούν να μετρηθούν στην αρχή της δοκιμής (όπως η αρτηριακή πίεση).

Οι αρχικές πληροφορίες παρουσιάζονται αποτελεσματικότερα σε έναν πίνακα (πίν. 4). Για συνεχείς μεταβλητές, όπως το βάρος ή η αρτηριακή πίεση, θα πρέπει να αναφέρεται η μεταβλητότητα των δεδομένων, μαζί με τις μέσες τιμές. Οι συνεχείς μεταβλητές μπορούν να συνοψιστούν για κάθε ομάδα με το μέσο όρο και την τυπική απόκλιση. Όταν τα συνεχή δεδομένα έχουν ασύμμετρη κατανομή, μια προτιμότερη προσέγγιση θα ήταν η παράθεση της διαμέ-

**Πίνακας 4.** Παράδειγμα αναφοράς βασικών δημογραφικών και κλινικών χαρακτηριστικών.\* (Προσαρμοσμένος από τον πίνακα 1 των Yusuf et al<sup>209</sup>).

	Τελμισαρτάνη (n=2954)	Εικονικό φάρμακο (n=2972)
Ηλικία (έτη)	66,9 (7,3)	66,9 (7,4)
Φύλο (θήλυ)	1.280 (43,3%)	1.267 (42,6%)
Καπνισματική κατάσταση		
Παρούσα	293 (9,9%)	289 (9,7%)
Παρελθούσα	1273 (43,1%)	1.283 (43,2%)
Εθνική καταγωγή		
Ασιάτες	637 (21,6%)	624 (21,0%)
Άραβες	37 (1,3%)	40 (1,3%)
Αφρικανοί	51 (1,7%)	55 (1,9%)
Ευρωπαίοι	1.801 (61,0%)	1.820 (61,2%)
Γηγενείς ή ιθαγενείς	390 (13,2%)	393 (13,2%)
Άλλο	38 (1,3%)	40 (1,3%)
Αρτηριακή πίεση (mmHg)	140,7 (16,8/81,8) (10,1)	141,3 (16,4/82,0) (10,2)
Καρδιακή συχνότητα (παλμοί/μίν)	68,8 (11,5)	68,8 (12,1)
Χοληστερόλη (mmol/L)		
Ολική	5,09 (1,18)	5,08 (1,15)
LDL	3,02 (1,01)	3,03 (1,02)
HDL	1,27 (0,37)	1,28 (0,41)
Στεφανιαία νόσος	2.211 (74,8%)	2.207 (74,3%)
Έμφραγμα μυοκαρδίου	1.381 (46,8%)	1.360 (45,8%)
Στηθάγγη	1.412 (47,8%)	1.412 (47,5%)
Περιφερική αγγειακή νόσος	349 (11,8%)	323 (10,9%)
Υπέρταση	2.259 (76,5%)	2.269 (76,3%)
Διαβήτης	1.059 (35,8%)	1.059 (35,6%)

\*Τα δεδομένα είναι μέσοι όροι (SD, τυπικές αποκλίσεις) ή αριθμοί (%)

σου και της εκατοστιαίας κλίμακας (όπως το 25ο και 75ο εκατοστημόριο).<sup>177</sup> Τα τυπικά σφάλματα και τα διαστήματα εμπιστοσύνης δεν είναι κατάλληλα για την περιγραφή της μεταβλητότητας αποτελούν επαγωγική παρά περιγραφική στατιστική. Μεταβλητές με μικρό αριθμό τακτικών κατηγοριών (όπως τα στάδια μιας νόσου από το I στο IV) δεν θα πρέπει να αντιμετωπίζονται ως συνεχείς μεταβλητές αντ' αυτού, θα πρέπει να αναφέρονται οι αριθμοί και τα ποσοστά για κάθε κατηγορία.<sup>48,177</sup>

Δυστυχώς, οι δοκιμασίες σημαντικότητας των αρχικών διαφορών είναι ακόμη συχνές.<sup>23,32,210</sup> Αναφέρονται στις μισές από τις 50 ΤΚΔ, οι οποίες δημοσιεύτηκαν σε κάποια από τα κορυφαία γενικά περιοδικά το 1997.<sup>183</sup> Τέτοιες δοκιμασίες σημαντικότητας εκτιμούν την πιθανότητα οι

παρατηρούμενες αρχικές διαφορές να έχουν προκύψει στην τύχη· εν τούτοις, γνωρίζουμε εκ των προτέρων ότι οι όποιες διαφορές οφείλονται στην τύχη. Οι δοκιμασίες των βασικών διαφορών δεν είναι απαραίτητα λανθασμένες, απλά είναι μη λογικές.<sup>211</sup> Παρόμοιος έλεγχος της υπόθεσης είναι περιττός και μπορεί να παραπλανήσει τους ερευνητές και τους αναγνώστες τους. Περισσότερο, οι συγκρίσεις των αρχικών βασικών χαρακτηριστικών θα πρέπει να βασίζονται στη θεώρηση της προγνωστικής δύναμης των μετρούμενων μεταβλητών και του μεγέθους πιθανών τυχαίων ανισοροπιών που προκύπτουν.<sup>211</sup>

**7.5. Στοιχείο 16.** Για κάθε ομάδα, ο αριθμός των συμμετεχόντων (παρονομαστής) που συμπεριελήφθησαν σε κάθε ανάλυση και κατά πόσο η ανάλυση έλαβε χώρα στις αρχικά ταξινομημένες ομάδες

*Παράδειγμα:* «Η πρωταρχική ανάλυση ήταν πρόθεση για θεραπεία και περιελάμβανε όλους τους ασθενείς, οι οποίοι ταξινομήθηκαν με τυχαίο τρόπο».<sup>212</sup>

«Ένας ασθενής στην ομάδα της αλενδρονάτης χάθηκε από την παρακολούθηση· έτσι, δεδομένα για την ανάλυση πρόθεσης για θεραπεία ήταν διαθέσιμα από 31 ασθενείς. Πέντε ασθενείς θεωρήθηκαν παραβάτες του πρωτοκόλλου... ως αποτέλεσμα 26 ασθενείς παρέμειναν για την ανάλυση, σύμφωνα με το πρωτόκολλο».<sup>213</sup>

*Εξήγηση:* Ο αριθμός των συμμετεχόντων σε κάθε ομάδα είναι ένα ουσιώδες στοιχείο της ανάλυσης. Αν και το διάγραμμα ροής (βλέπε στοιχείο 13α) μπορεί να υποδεικνύει τον αριθμό των συμμετεχόντων που αναλύθηκαν, οι αριθμοί αυτοί συχνά ποικίλλουν στα διάφορα μέτρα των εκβάσεων. Ο αριθμός των συμμετεχόντων ανά ομάδα θα πρέπει να δίνεται για όλες τις αναλύσεις. Για διχότομες εκβάσεις (όπως ο λόγος κινδύνου ή η διαφορά κινδύνου), οι παρονομαστές ή τα ποσοστά εκδήλωσης θα πρέπει να αναφέρονται. Η έκφραση των αποτελεσμάτων ως ποσοστά επίσης βοηθά τον αναγνώστη να εκτιμήσει αν κάποιοι από τους τυχαία ταξινομημένους συμμετέχοντες αποκλείστηκαν από την ανάλυση. Επομένως, τα αποτελέσματα δεν θα πρέπει να αναφέρονται μόνο ως συνοπτικά μέτρα, όπως ως σχετικός κίνδυνος.

Οι συμμετέχοντες, μερικές φορές, ενδέχεται να μη λάβουν την πλήρη παρέμβαση ή μερικοί μη κατάλληλοι ασθενείς μπορεί να διανεμήθηκαν με τυχαίο τρόπο εσφαλμένα. Ένας ευρέως προτεινόμενος τρόπος αντιμετώπισης τέτοιων θεμάτων είναι η ανάλυση όλων των συμμετεχόντων σύμφωνα με την αρχική τους ταξινόμηση σε ομάδες, ανεξάρτητα από το τι συνέβη στη συνέχεια (βλέπε πλαίσιο 6). Αυτή η

στρατηγική «πρόθεση για θεραπεία» δεν είναι πάντα σαφής στην εφαρμογή της. Είναι σύνηθες για μερικούς ασθενείς να μην ολοκληρώνουν μια μελέτη –μπορεί να εγκαταλείψουν ή να αποσυρθούν από την ενεργό θεραπεία– και έτσι να μην εκτιμηθούν στο τέλος. Εάν η έκβαση είναι η θνησιμότητα, τέτοιοι ασθενείς μπορεί να συμπεριληφθούν στην ανάλυση, σύμφωνα με πληροφορίες αρχείου, ενώ ενδέχεται να χρειαστεί να χρησιμοποιηθούν τεχνικές αντικατάστασης (imputation techniques), αν ελλείπουν άλλα δεδομένα έκβασης. Ο όρος «ανάλυση πρόθεσης για θεραπεία» συχνά χρησιμοποιείται με ακατάλληλο τρόπο για παράδειγμα, όταν αυτοί οι οποίοι δεν έλαβαν την πρώτη δόση ενός φαρμάκου υπό δοκιμή αποκλείονται από την ανάλυση.<sup>18</sup>

Αντίθετα, η ανάλυση μπορεί να περιοριστεί μόνο στους συμμετέχοντες οι οποίοι εκπληρώνουν το πρωτόκολλο όσον αφορά στην καταλληλότητα, στις παρεμβάσεις και στην εκτίμηση της έκβασης. Αυτή η ανάλυση είναι γνωστή ως «ανάλυση σε θεραπεία» ή «ανάλυση σύμφωνα με το πρωτόκολλο». Ο αποκλεισμός των συμμετεχόντων από την ανάλυση μπορεί να οδηγήσει σε εσφαλμένα συμπεράσματα. Για παράδειγμα, σε μια δοκιμή που συνέκρινε την ιατρική με τη χειρουργική θεραπεία για τη στένωση καρωτίδων, η ανάλυση η οποία περιορίστηκε στους συμμετέχοντες που ήταν διαθέσιμοι για επανεξέταση έδειξε ότι η χειρουργική αντιμετώπιση μείωσε τον κίνδυνο για παροδικά ισχαιμικά επεισόδια, αγγειακό εγκεφαλικό επεισόδιο και θάνατο. Εν τούτοις, η ανάλυση πρόθεσης για θεραπεία η οποία βασίστηκε στην αρχική ταξινόμηση όλων των συμμετεχόντων δεν έδειξε ανώτερο αποτέλεσμα της χειρουργικής αντιμετώπισης.<sup>214</sup>

Η ανάλυση πρόθεσης για θεραπεία γενικά προτιμάται, καθώς αποφεύγει σφάλματα σχετιζόμενα με μη τυχαία απώλεια συμμετεχόντων.<sup>215–217</sup> Ανεξάρτητα από το αν οι συγγραφείς χρησιμοποιούν τον όρο «πρόθεση για θεραπεία», θα πρέπει να καταστήσουν σαφές ποιοι και πόσοι συμμετέχοντες περιλαμβάνονται σε κάθε ανάλυση (βλέπε στοιχείο 13). Μη συμμόρφωση με την ανατεθείσα θεραπεία μπορεί να σημαίνει ότι η ανάλυση πρόθεσης για θεραπεία υποεκτιμά το πιθανό όφελος της θεραπείας και ίσως γι' αυτό θα πρέπει να τεθεί υπό σκέψη η εφαρμογή συμπληρωματικής ανάλυσης, όπως η ανάλυση σύμφωνα με το πρωτόκολλο.<sup>218,219</sup> Πρέπει να σημειωθεί, εν τούτοις, ότι τέτοιες αναλύσεις είναι συχνά σημαντικά εσφαλμένες.<sup>220</sup>

Σε μια ανασκόπηση 403 ΤΚΔ που δημοσιεύτηκαν σε 10 κορυφαία ιατρικά περιοδικά το 2002, οι 249 (63%) ανέφεραν τη χρήση ανάλυσης πρόθεσης για θεραπεία ως την κύρια ανάλυση. Η αναλογία αυτή ήταν υψηλότερη για περιοδικά τα οποία υιοθετούν τη δήλωση CONSORT

(70% έναντι 48%). Από τα άρθρα τα οποία ανέφεραν τη χρήση ανάλυσης πρόθεσης για θεραπεία, μόνο στο 39% πραγματικά αναλύθηκαν οι συμμετέχοντες, όπως τυχαιοποιήθηκαν, με περισσότερα από 60% των άρθρων να έχουν ελλιπή δεδομένα στην κύρια ανάλυση.<sup>221</sup> Άλλες μελέτες δείχνουν παρόμοια ευρήματα.<sup>18,222,223</sup> Οι δοκιμές οι οποίες δεν αναφέρουν αποκλεισμούς είναι κατά τα άλλα μεθοδολογικά ασθενέστερες από εκείνες που αναφέρουν σχετικά με τον αποκλεισμό συμμετεχόντων,<sup>173</sup> γεγονός που παρέχει ισχυρές ενδείξεις ότι τουλάχιστον κάποιοι ερευνητές οι οποίοι απέκλεισαν συμμετέχοντες δεν το ανέφεραν. Άλλη μελέτη βρήκε ότι η αναφορά ανάλυσης πρόθεσης για θεραπεία σχετιζόταν και με άλλα στοιχεία καλού σχεδιασμού και αναφοράς της μελέτης, όπως η περιγραφή του υπολογισμού του μεγέθους του δείγματος.<sup>224</sup>

7.6. Στοιχείο 17α. Για κάθε κύρια και δευτερεύουσα έκβαση, αποτελέσματα για κάθε ομάδα, και το εκτιμώμενο μέγεθος επίδρασης, καθώς και η ακρίβειά του (όπως τα 95% διαστήματα εμπιστοσύνης)

*Παράδειγμα:* Βλέπε πίνακες 5 και 6.

*Εξήγηση:* Για κάθε έκβαση, τα αποτελέσματα της μελέτης θα πρέπει να αναφέρονται ως μια περίληψη των εκβάσεων σε κάθε ομάδα (για παράδειγμα, ο αριθμός των συμμετεχόντων με ή χωρίς το συμβάν και οι παρονομαστές ή η μέση τιμή και η τυπική απόκλιση των μετρήσεων), μαζί με τη σύγκριση μεταξύ των ομάδων, γνωστή και ως μέγεθος επίδρασης. Για διχοτόμες εκβάσεις, το μέγεθος επίδρασης μπορεί να είναι ο λόγος κινδύνου (σχετικός κίνδυνος), ο λόγος αναλογιών ή η διαφορά κινδύνου<sup>7</sup> για δεδομένα του χρόνου επιβίωσης, μπορεί να είναι η αναλογία κινδύνων ή η διαφορά στο διάμεσο χρόνο επιβίωσης, ενώ για συνεχή δεδομένα είναι συνήθως η διαφορά των μέσων τιμών. Τα διαστήματα εμπιστοσύνης θα πρέπει να παρουσιάζονται για τη σύγκριση μεταξύ των ομάδων. Ένα σύνηθες λάθος είναι η παρουσίαση ξεχωριστών διαστημάτων εμπιστοσύνης για την έκβαση σε κάθε ομάδα, παρά για το θεραπευτικό αποτέλεσμα.<sup>233</sup> Τα αποτελέσματα μιας δοκιμής παρουσιάζονται συχνά ευκρινέστερα σε έναν πίνακα παρά στο κείμενο, όπως φαίνεται στους πίνακες 5 και 6.

Για όλες τις εκβάσεις, οι συγγραφείς πρέπει να παρέχουν ένα διάστημα εμπιστοσύνης που υποδεικνύει την ακρίβεια (αβεβαιότητα) της εκτίμησης.<sup>48,235</sup> Το 95% διάστημα εμπιστοσύνης είναι συμβατικό, αλλά περιστασιακά χρησιμοποιούνται και άλλα επίπεδα. Πολλά περιοδικά απαιτούν ή ενθαρρύνουν έντονα τη χρήση των διαστημάτων εμπιστοσύνης.<sup>236</sup> Είναι ιδιαίτερα πολύτιμα σε σχέση με διαφορές οι οποίες δεν



**Πίνακας 5.** Παράδειγμα αναφοράς συνοπτικών αποτελεσμάτων για κάθε ομάδα της μελέτης (διχότομες εκβάσεις).\* (Προσαρμοσμένος από τον πίνακα 2 των Mease et al<sup>103</sup>).

Καταληκτικό σημείο	Αριθμός (%)		Διαφορά κινδύνου (95% ΔΕ)
	Etanercept (n=30)	Εικονικό φάρμακο (n=30)	
<i>Κύριο καταληκτικό σημείο</i>			
Επίτευξη PsARC σε 12 εβδομάδες	26 (87)	7 (23)	63% (44–83)
<i>Δευτερεύον καταληκτικό σημείο</i>			
Ποσοστό των ασθενών που πληρούν τα κριτήρια ACR			
ACR20	22 (73)	4 (13)	60% (40–80)
ACR50	15 (50)	1 (3)	47% (28–66)
ACR70	4 (13)	0 (0)	13% (1–26)

\*Ανατρέξτε επίσης στο παράδειγμα για στοιχείο 6α.

PsARC: Psoriasis Arthritis Response Criteria (κριτήρια ανταπόκρισης ψωριασικής αρθρίτιδας), ACR: American College of Rheumatology (Αμερικανικό Κολέγιο Ρευματολογίας)

**Πίνακας 6.** Παράδειγμα αναφοράς συνοπτικών αποτελεσμάτων για κάθε ομάδα της μελέτης (συνεχείς εκβάσεις). (Προσαρμοσμένος από τον πίνακα 3 του Van Linschoten<sup>234</sup>).

	Θεραπεία άσκησης (n=65)		Ομάδα ελέγχου (n=66)		Σταθμισμένη διαφορά* (95% ΔΕ) στους 12 μήνες
	Έναρξη (μέση τιμή, SD)	12 μήνες (μέση τιμή, SD)	Έναρξη (μέση τιμή, SD)	12 μήνες (μέση τιμή, SD)	
Δείκτης λειτουργίας (0–100)	66,4 (13,9)	83,2 (14,8)	65,9 (15,2)	79,8 (17,5)	4,52 (-0,73–9,76)
Πόνος κατά την ανάπαυση (0–100)	4,14 (2,3)	1,43 (2,2)	4,03 (2,3)	2,61 (2,9)	-1,29 (-2,16–0,42)
Πόνος κατά τη δραστηριότητα (0–100)	6,32 (2,2)	2,57 (2,9)	5,97 (2,3)	3,54 (3,38)	-1,19 (-2,22 έως -0,16)

\*Ο δείκτης λειτουργίας σταθμισμένος για την ηλικία έναρξης και τη διάρκεια των συμπτωμάτων

φθάνουν σε επίπεδα συμβατικής στατιστικής σημαντικότητας, υποδεικνύοντας ότι τα αποτελέσματα δεν αποκλείουν μια σημαντική κλινική διαφορά. Η χρήση των διαστημάτων εμπιστοσύνης έχει αυξηθεί σημαντικά τα τελευταία χρόνια, αν και όχι σε όλες τις ιατρικές ειδικότητες.<sup>233</sup> Αν και τιμές του  $p$  παρέχονται μαζί με τα διαστήματα εμπιστοσύνης, τα αποτελέσματα δεν θα πρέπει να αναφέρονται μόνο ως τιμές  $p$ .<sup>237,238</sup> Αποτελέσματα θα πρέπει να αναφέρονται για όλα τα σχεδιαζόμενα κύρια και δευτερεύοντα καταληκτικά σημεία και όχι μόνο για αναλύσεις οι οποίες ήταν στατιστικά σημαντικές ή «ενδιαφέρουσες». Η επιλεκτική αναφορά σε μια μελέτη είναι ένα διαδεδομένο και σοβαρό πρόβλημα.<sup>55,57</sup> Σε δοκιμές στις οποίες διενεργήθησαν ενδιάμεσες αναλύσεις, η ερμηνεία θα πρέπει να επικεντρώνεται στα τελικά αποτελέσματα κατά το πέρας της δοκιμής και όχι στα ενδιάμεσα αποτελέσματα.<sup>239</sup>

Τόσο για τα διχότομα δεδομένα όσο και για τα δεδομένα του χρόνου επιβίωσης μπορεί να είναι χρήσιμη η έκφραση των αποτελεσμάτων ως ο αριθμός των ασθενών που απαιτείται να λάβουν θεραπεία για να ωφεληθεί ή να παρουσιάσει βλάβη ένας ασθενής (βλέπε στοιχείο 21).

**7.7. Στοιχείο 17β.** Για διχότομες εκβάσεις, συστήνεται η παρουσίαση τόσο των απόλυτων, όσο και των σχετικών μεγεθών επίδρασης

*Παράδειγμα:* «Ο κίνδυνος εξάρτησης από το οξυγόνο ή ο θάνατος μειώθηκε κατά 16% (95% ΔΕ 25–7%). Η απόλυτη διαφορά ήταν -6,3% (95% ΔΕ -9,9% έως -2,7%). Συνεπώς, η έγκαιρη χορήγηση σε έναν υπολογιζόμενο αριθμό 16 βρεφών θα είχε αποτρέψει ένα βρεφικό θάνατο ή τη μακρόχρονη εξάρτηση από το οξυγόνο» (ακόμη βλέπε πίν. 7).<sup>242</sup>

*Εξήγηση:* Όταν η κύρια έκβαση είναι διχότομη, η σχετική επίδραση (λόγος κινδύνου [σχετικός κίνδυνος] ή λόγος αναλογιών) και η απόλυτη επίδραση (διαφορά κινδύνου) πρέπει να αναφέρονται (με διαστήματα εμπιστοσύνης), αφού ούτε μόνο τα σχετικά μέτρα ούτε μόνο τα απόλυτα μέτρα δεν παρέχουν ολοκληρωμένη εικόνα της επίδρασης και των επιπτώσεών της. Διαφορετικά ακροατήρια μπορεί να προτιμούν είτε το σχετικό είτε τον απόλυτο κίνδυνο, αλλά τόσο οι ιατροί όσο και ο κοινός λαός τείνουν να υπερεκτιμούν την επίδραση, όταν αυτή παρουσιάζεται με τους όρους του σχετικού κινδύνου.<sup>243–245</sup> Το μέγεθος της διαφοράς κινδύνου είναι λιγότερο γενικεύσιμο σε άλλους

**Πίνακας 7.** Παράδειγμα αναφοράς απόλυτου και σχετικού μεγέθους επίδρασης (προσαρμοσμένος από τον πίνακα 3 της ομάδας συνεργασίας OSIRIS<sup>242</sup>).

Κύρια έκβαση	Ποσοστά (no)		Λόγος κινδύνου (95% ΔΕ)	Διαφορά κινδύνου (95% ΔΕ)
	Πρώιμη χορήγηση (n=1.344)	Καθυστερημένη επιλεκτική χορήγηση (n=1.346)		
Θάνατος ή εξάρτηση από οξυγόνο «σε πιθανή ημερομηνία τοκετού»	31,9 (429)	38,2 (514)	0,84 (0,75 – 0,93)	-6,3 (-9,9 έως -2,7)

πληθυσμούς από το σχετικό κίνδυνο, μιας και εξαρτάται από τον αρχικό κίνδυνο στην ομάδα των ατόμων που δεν εκτίθενται, ο οποίος αρχικός κίνδυνος ποικίλλει στους διάφορους πληθυσμούς. Για νόσους των οποίων η έκβαση είναι συνήθης, ένας σχετικός κίνδυνος πλησίον της μονάδας μπορεί να υποδεικνύει κλινικά σημαντικές διαφορές από την άποψη της δημόσιας υγείας. Σε αντίθεση, ένας μεγάλος σχετικός κίνδυνος, όταν η έκβαση είναι σπάνια, μπορεί να μην είναι τόσο σημαντικός για τη δημόσια υγεία (παρ' όλο που μπορεί να είναι σημαντικός για ένα άτομο σε κατηγορία υψηλού κινδύνου).

#### 7.8. Στοιχείο 18. Αποτελέσματα κάθε άλλης ανάλυσης, περιλαμβανομένης της ανάλυσης υποομάδων και σταθμισμένης ανάλυσης, διακρίνοντας την προκαθορισμένη από τη διερευνητική

*Παράδειγμα:* «Στη βάση μιας μελέτης η οποία προτείνει ότι η περιεχειρητική αποτελεσματικότητα των β-αποκλειστών μπορεί να ποικίλλει ανάλογα με τον αρχικό κίνδυνο, προκαθορίσαμε την κύρια ανάλυση υποομάδων με βάση το αναθεωρημένο σύστημα βαθμολόγησης του δείκτη του καρδιακού κινδύνου. Επίσης, προκαθορίσαμε τη δευτερεύουσα ανάλυση υποομάδων με βάση το φύλο, το είδος του χειρουργείου και τη χρήση επισκληρίδιας ή ραχιαίας αναισθησίας. Για όλες τις αναλύσεις των υποομάδων χρησιμοποιήσαμε το μοντέλο αναλογικού κινδύνου του Cox, το οποίο ενσωμάτωνε ελέγχους για αλληλεπιδράσεις που ήταν σχεδιασμένες να είναι σημαντικές σε επίπεδο  $p < 0,05$ ... Η εικόνα 3 δείχνει τα αποτελέσματα των προκαθορισμένων αναλύσεων των υποομάδων και υποδεικνύει συνέπεια των επιδράσεων... Οι αναλύσεις των υποομάδων μας είχαν χαμηλή ισχύ στην ανίχνευση μέτριων διαφορών στα αποτελέσματα των υποομάδων που κάποιος θα περίμενε να ανιχνευτούν, εάν υπήρχε πραγματική επίδραση υποομάδων<sup>100</sup>».

*Εξήγηση:* Πολλαπλές αναλύσεις των ίδιων δεδομένων δημιουργούν τον κίνδυνο για ψευδώς θετικά ευρήματα.<sup>246</sup> Οι συγγραφείς θα πρέπει να αντιστέκονται στον πειρασμό για πολλές αναλύσεις υποομάδων.<sup>183,185,247</sup> Οι αναλύσεις οι οποίες είναι προκαθορισμένες στο πρωτόκολλο της δοκιμής (βλέπε στοιχείο 24) είναι πολύ περισσότερο αξιόπιστες

από αυτές που προκύπτουν από τα δεδομένα, και γι' αυτό οι συγγραφείς θα πρέπει να αναφέρουν ποιες αναλύσεις είχαν προκαθοριστεί. Εάν πραγματοποιήθηκαν αναλύσεις υποομάδων, οι συγγραφείς θα πρέπει να αναφέρουν ποιες υποομάδες ελέγχθηκαν, γιατί, εάν ήταν προκαθορισμένες, και πόσες ήταν προκαθορισμένες. Η επιλεκτική αναφορά των αναλύσεων των υποομάδων μπορεί να οδηγήσει σε σφάλματα.<sup>248</sup> Όταν εκτιμάται μια υποομάδα, το ερώτημα δεν είναι αν η υποομάδα εμφάνισε ένα αποτέλεσμα με στατιστική σημαντικότητα, αλλά εάν το θεραπευτικό αποτέλεσμα στην υποομάδα ήταν στατιστικά διαφορετικό από τις άλλες. Για να καθοριστεί αυτό, ένας έλεγχος των αλληλεπιδράσεων είναι επικουρικός, παρ' όλο που η ισχύς τέτοιων δοκιμασιών είναι τυπικά χαμηλή. Εάν πραγματοποιήθηκαν επίσημες αξιολογήσεις των αλληλεπιδράσεων (βλέπε στοιχείο 12β), θα πρέπει να αναφέρονται ως εκτιμώμενη διαφορά στο αποτέλεσμα της παρέμβασης σε κάθε υποομάδα (με επίπεδο εμπιστοσύνης) και όχι μόνο ως τιμές  $p$ .

Σε μια έρευνα, 35 από τις 50 εργασίες δοκιμών περιλάμβαναν αναλύσεις υποομάδων, από τις οποίες μόνο σε ποσοστό 42% χρησιμοποιήθηκαν έλεγχοι αλληλεπίδρασης.<sup>183</sup> Συχνά, ήταν δύσκολο να καθοριστεί εάν οι αναλύσεις των υποομάδων είχαν καθοριστεί στο πρωτόκολλο. Σε μια άλλη έρευνα χειρουργικών δοκιμών, οι οποίες δημοσιεύτηκαν σε περιοδικά με υψηλό συντελεστή απήχησης, 27 στις 72 δοκιμές ανέφεραν 54 αναλύσεις υποομάδων, από τις οποίες 92% ήταν εκ των υστέρων και μόνο 6% των αναλύσεων υποομάδων χρησιμοποίησαν έλεγχο αλληλεπίδρασης για την εκτίμηση της ύπαρξης μιας επίδρασης υποομάδας.<sup>249</sup>

Παρόμοιες συστάσεις ισχύουν για αναλύσεις, στις οποίες η προσαρμογή έγινε για τις βασικές μεταβλητές. Εάν έγινε τέτοια ανάλυση, θα πρέπει να αναφερθεί τόσο αυτή, όσο και οι αναλύσεις χωρίς προσαρμογή. Οι συγγραφείς πρέπει να δηλώνουν εάν σχεδιάστηκαν σταθμισμένες αναλύσεις, περιλαμβάνοντας την επιλογή των μεταβλητών για τις οποίες έγινε στάθμιση. Ιδανικά, στο πρωτόκολλο της δοκιμής θα πρέπει να δηλώνεται εάν έγινε στάθμιση σε οριζόμενες βασικές αρχικές μεταβλητές χρησιμοποιώντας ανάλυση συνδιακύμανσης.<sup>187</sup> Η προσαρμογή για μεταβλητές λόγω της σημαντικής διαφοράς τους κατά την έναρξη είναι πιθανό να οδηγήσει σε σφάλμα στην

εκτίμηση του θεραπευτικού αποτελέσματος.<sup>187</sup> Μια έρευνα διαπίστωσε ότι μη αναγνωρισμένες αποκλίσεις μεταξύ των πρωτοκόλλων και των δημοσιεύσεων βρέθηκαν και στις 25 δοκιμές στις οποίες αναφέρθηκαν αναλύσεις υποομάδων, και στις 23 από τις 28 δοκιμές στις οποίες αναφέρθηκαν σταθμισμένες αναλύσεις.<sup>92</sup>

### 7.9. Στοιχείο 19. Όλες οι σημαντικές βλαβερές συνέπειες ή οι ακούσιες επιδράσεις σε κάθε ομάδα

Για ειδική καθοδήγηση, ανατρέξτε στο CONSORT όσον αφορά στις βλαβερές συνέπειες.<sup>42</sup>

*Παράδειγμα:* «Το ποσοστό των ασθενών που βίωσαν κάποια ανεπιθύμητη ενέργεια ήταν παρόμοια μεταξύ των ομάδων του rBPI21 (ανασυνδυασμένη βακτηριοκτόνος/αυξάνουσα τη διαπερατότητα πρωτεΐνης, recombinant bactericidal/permeability-increasing protein) και του εικονικού φαρμάκου: 168 (88,4%) από τους 190 και 180 (88,7%) από τους 203, αντίστοιχα, και ήταν λιγότερες στους ασθενείς, οι οποίοι αντιμετωπίστηκαν με rBPI21 έναντι εκείνων που αντιμετωπίστηκαν με το εικονικό φάρμακο σε 11 από τα 12 συστήματα του σώματος... το ποσοστό των ασθενών οι οποίοι βίωσαν μια σοβαρή ανεπιθύμητη ενέργεια, όπως κρίθηκε από τους ερευνητές, ήταν αριθμητικά μικρότερο στην ομάδα του rBPI21 έναντι της ομάδας του εικονικού φαρμάκου: 53 (27,9%) στους 190 έναντι 74 (36,5%) στους 203 ασθενείς, αντίστοιχα. Αναφέρθηκαν ως σχετικές με φάρμακα μόνο 3 περιπτώσεις σοβαρών ανεπιθύμητων ενεργειών και όλες συνέβησαν στην ομάδα του εικονικού φαρμάκου.<sup>250</sup>

*Εξήγηση:* Οι αναγνώστες χρειάζονται πληροφορίες σχετικά με τις βλαβερές συνέπειες, καθώς και τα οφέλη των παρεμβάσεων για να λάβουν λογικές και ισορροπημένες αποφάσεις. Η ύπαρξη και η φύση των ανεπιθύμητων ενεργειών μπορούν να έχουν σημαντική επίπτωση στο αν μια συγκεκριμένη παρέμβαση θα θεωρηθεί αποδεκτή και χρήσιμη. Όλες οι αναφερόμενες ανεπιθύμητες ενέργειες που παρατηρούνται κατά τη διάρκεια μιας δοκιμής δεν είναι απαραίτητα συνέπεια της παρέμβασης· μερικές μπορεί να είναι συνέπεια της κατάστασης η οποία αντιμετωπίζεται. Οι τυχαίοποιημένες δοκιμές προσφέρουν την καλύτερη προσέγγιση για παροχή τόσο δεδομένων ασφάλειας όσο και δεδομένων αποτελεσματικότητας, αν και δεν μπορούν να ανιχνεύσουν σπάνιες βλαβερές συνέπειες.

Πολλές αναφορές των ΤΚΔ παρέχουν ανεπαρκείς πληροφορίες σχετικά με τις ανεπιθύμητες ενέργειες. Μια έρευνα 192 δοκιμών φαρμάκων που δημοσιεύτηκαν από το 1967–1999 έδειξε ότι μόνο σε ποσοστό 39% υπήρχε επαρκής αναφορά των κλινικών ανεπιθύμητων ενεργειών

και σε 29% υπήρχε επαρκής αναφορά της εργαστηριακά οριζόμενης τοξικότητας.<sup>72</sup> Πιο πρόσφατα, μια σύγκριση μεταξύ των δεδομένων των ανεπιθύμητων ενεργειών, τα οποία υποβάλλονται στη βάση δεδομένων των δοκιμών του Εθνικού Ινστιτούτου για τον Καρκίνο, το οποίο χρηματοδότησε τις δοκιμές, και των πληροφοριών που αναφέρονται σε άρθρα περιοδικών βρήκε ότι οι ανεπιθύμητες ενέργειες χαμηλού βαθμού υποαναφέρονται στα άρθρα των περιοδικών. Συμβάντα υψηλού βαθμού (βαθμοί κοινών κριτηρίων τοξικότητας 3–5) αναφέρονταν ασυνεπώς στα άρθρα, και οι πληροφορίες σχετικά με την απόδοσή τους στα υπό έρευνα φάρμακα ήταν ελλιπείς.<sup>251</sup> Επί πλέον, μια ανασκόπηση των δοκιμών οι οποίες δημοσιεύτηκαν σε έξι γενικά ιατρικά περιοδικά από το 2006–2007 βρήκε ότι, παρ' όλο που 89% από τις 133 αναφορές ανέφεραν ανεπιθύμητες ενέργειες, δεν δόθηκε κάποια πληροφορία για σοβαρές ανεπιθύμητες ενέργειες και για αποχώρηση των ασθενών εξ αιτίας μιας ανεπιθύμητης ενέργειας σε ποσοστό 27% και 48% των άρθρων, αντίστοιχα.<sup>252</sup>

Μια επέκταση της δήλωσης CONSORT έχει δημιουργηθεί για να παρέχει λεπτομερείς συστάσεις σχετικά με την αναφορά των βλαβερών συνεπειών στις τυχαίοποιημένες δοκιμές.<sup>42</sup> Συστάσεις και παραδείγματα κατάλληλης αναφοράς είναι διαθέσιμα ελεύθερα στην ιστοσελίδα της δήλωσης CONSORT 2010 ([www.consort-statement.org](http://www.consort-statement.org)) και θα πρέπει να τα συμβουλευτεί ο καθένας, ιδιαίτερα εάν η μελέτη των βλαβερών συνεπειών αποτελεί έναν κύριο στόχο. Εν συντομία, εάν έχουν συλλεχθεί δεδομένα σχετικά με τις ανεπιθύμητες ενέργειες, τα συμβάματα θα πρέπει να αριθμούνται και να ορίζονται, με αναφορά σε τυποποιημένα κριτήρια, όπου αυτό είναι κατάλληλο. Οι μέθοδοι που χρησιμοποιήθηκαν για τη συλλογή δεδομένων και την απόδοση των συμβαμάτων θα πρέπει να περιγράφονται. Για κάθε σκέλος της μελέτης θα πρέπει να παρουσιάζεται ο απόλυτος κίνδυνος κάθε ανεπιθύμητης ενέργειας, χρησιμοποιώντας τις κατάλληλες μετρήσεις για επαναλαμβανόμενα συμβάματα, καθώς και ο αριθμός των συμμετεχόντων που αποχώρησαν λόγω βλαβερών συνεπειών. Τελειώνοντας, οι συγγραφείς θα πρέπει να παρέχουν μια ισορροπημένη συζήτηση των ωφελειών και των βλαβερών συνεπειών.<sup>42</sup>

## 8. ΣΥΖΗΤΗΣΗ

### 8.1. Στοιχείο 20. Περιορισμοί της δοκιμής, αντιμετώπιζοντας τις πηγές πιθανών σφαλμάτων, την ανακρίβεια, και αν είναι σχετικό, την πολλαπλότητα των αναλύσεων

*Παράδειγμα:* «Η υπεροχή των αρρένων ασθενών (85%) είναι ένας περιορισμός της μελέτης μας... Χρησιμοποιή-



σαμε μη επικαλυμμένους μεταλλικούς νάρθηκες (stents), αφού οι νάρθηκες που απελευθερώνουν φάρμακο δεν ήταν διαθέσιμοι έως αργά, κατά τη διάρκεια της εγγραφής ασθενών. Αν και ο τελευταίος παράγοντας μπορεί να εκληφθεί ως περιορισμός, δημοσιευμένα δεδομένα δεν υποδεικνύουν κάποιο όφελος (είτε βραχυπρόθεσμο, είτε μακροπρόθεσμο) σε σχέση με το θάνατο και το έμφραγμα του μυοκαρδίου σε ασθενείς με σταθερή στεφανιαία νόσο οι οποίοι έλαβαν νάρθηκες που απελευθερώνουν φάρμακο, σε σύγκριση με αυτούς που έλαβαν μη επικαλυμμένους μεταλλικούς νάρθηκες».<sup>253</sup>

*Εξήγηση:* Το τμήμα της συζήτησης των επιστημονικών δημοσιεύσεων συχνά πληρούται από ρητορική υποστήριξη των ευρημάτων των συγγραφέων<sup>254</sup> και παρέχει λίγα μετρούμενα επιχειρήματα των υπέρ και των κατά της μελέτης και των αποτελεσμάτων της. Κάποια περιοδικά έχουν προσπαθήσει να διορθώσουν αυτό το πρόβλημα, ενθαρρύνοντας μεγαλύτερη δομή στη συζήτηση των αποτελεσμάτων από τους συγγραφείς.<sup>255,256</sup> Για παράδειγμα, το περιοδικό *Annals of Internal Medicine* προτείνει στους συγγραφείς να δομήσουν το τμήμα της συζήτησης παρουσιάζοντας (α) μια βραχεία σύνοψη των κύριων ευρημάτων, (β) μια θεώρηση των πιθανών μηχανισμών και εξηγήσεων, (γ) σύγκριση με σχετικά ευρήματα από άλλες δημοσιευμένες μελέτες (όπου αυτό είναι δυνατό, περιλαμβάνοντας μια συστηματική ανασκόπηση που θα συνδυάζει τα αποτελέσματα της παρούσας μελέτης με αυτά όλων των προηγούμενων σχετικών μελετών), (δ) τους περιορισμούς της συγκεκριμένης μελέτης (και μεθόδους που χρησιμοποιήθηκαν για την ελαχιστοποίηση ή την αντιστάθμιση αυτών των περιορισμών) και (ε) ένα βραχύ τμήμα, το οποίο συνοψίζει τις κλινικές και τις ερευνητικές επιπτώσεις της εργασίας, ανάλογα με την περίπτωση.<sup>255</sup> Προτείνουμε στους συγγραφείς να ακολουθούν τις λογικές αυτές προτάσεις, πιθανόν κάνοντας χρήση των κατάλληλων υπότιτλων στο τμήμα της συζήτησης.

Παρόλο που η συζήτηση των περιορισμών συχνά παραλείπεται από τις ερευνητικές αναφορές,<sup>257</sup> η αναγνώριση και η συζήτηση των αδυναμιών μιας μελέτης έχει ιδιαίτερη σημασία.<sup>258</sup> Για παράδειγμα, μια χειρουργική ομάδα ανέφερε ότι η λαπαροσκοπική χολοκυστεκτομή, μια τεχνικά δύσκολη διαδικασία, είχε σημαντικά χαμηλότερα ποσοστά επιπλοκών από την πιο κλασική, ανοικτή χολοκυστεκτομή για την αντιμετώπιση της οξείας χολοκυστίτιδας.<sup>259</sup> Εν τούτοις, οι συγγραφείς παρέλειψαν να συζητήσουν ένα εμφανές σφάλμα στα αποτελέσματά τους. Όλες οι λαπαροσκοπικές χολοκυστεκτομές είχαν ολοκληρωθεί από τους ερευνητές της μελέτης, ενώ 80% των ανοικτών χολοκυστεκτομών είχαν ολοκληρωθεί από εκπαιδευμένους.

Οι συγγραφείς θα πρέπει ακόμη να συζητούν πιθανές

ανακρίβειες των αποτελεσμάτων. Ανακρίβειες μπορεί να προκύψουν σε διάφορες πτυχές της μελέτης, περιλαμβανομένης της μέτρησης της κύριας έκβασης (βλέπε στοιχείο 6α) ή της διάγνωσης (βλέπε στοιχείο 4α). Ίσως η κλίμακα που χρησιμοποιήθηκε να ήταν σταθμισμένη σε πληθυσμό ενηλίκων και να χρησιμοποιήθηκε σε παιδιατρικό πληθυσμό ή ο εκτιμητής δεν ήταν εκπαιδευμένος στη διαχείριση του οργάνου.

Πάντα θα πρέπει να λαμβάνεται υπ' όψη η διαφορά μεταξύ της στατιστικής σημαντικότητας και της κλινικής σημασίας. Οι συγγραφείς θα πρέπει ιδιαίτερα να αποφεύγουν το σύνθημα λάθος της ερμηνείας ενός μη στατιστικά σημαντικού αποτελέσματος ως ενδεικτικού ισοδυναμίας των παρεμβάσεων. Τα διαστήματα εμπιστοσύνης (βλέπε στοιχείο 17α) παρέχουν πολύτιμες πληροφορίες σχετικά με το αν το αποτέλεσμα της δοκιμής είναι συμβατό με μια σημαντική κλινικά επίδραση, ανεξάρτητα από την τιμή p.<sup>120</sup>

Οι συγγραφείς θα πρέπει να επιδεικνύουν ιδιαίτερη φροντίδα, όταν αξιολογούν τα αποτελέσματα δοκιμών με πολλαπλές συγκρίσεις. Τέτοια πολλαπλότητα προκύπτει από αρκετές παρεμβάσεις, μέτρα έκβασης, χρονικά σημεία, αναλύσεις υποομάδων και άλλους παράγοντες. Υπό τέτοιες συνθήκες, κάποια στατιστικώς σημαντικά ευρήματα ενδέχεται να έχουν προκύψει μόνο από τύχη.

## 8.2. Στοιχείο 21. Γενικευσιμότητα (εξωτερική εγκυρότητα, εφαρμοσιμότητα) των ευρημάτων της δοκιμής

*Παραδείγματα:* «Καθώς η παρέμβαση υλοποιήθηκε και για τα δύο φύλα, για όλες τις ηλικίες, όλα τα είδη των αθλημάτων και τα διαφορετικά επίπεδα των αθλημάτων, τα αποτελέσματα υποδεικνύουν ότι ολόκληρο το εύρος των αθλητών, από την αφρόκρεμα των νεαρών αθλητών έως τους ενδιάμεσους και πρεσβύτερους αθλητές αναψυχής, θα ωφεληθεί από τη χρήση του συγκεκριμένου προγράμματος για την πρόληψη υποτροπών του διαστρέμματος του αστραγάλου. Συμπεριλαμβάνοντας τους αθλητές οι οποίοι θεραπεύτηκαν μη ιατρικά και ιατρικά, καλύψαμε ένα ευρύ φάσμα της σοβαρότητας της κάκωσης. Αυτό υποδηλώνει ότι το παρόν πρόγραμμα προπόνησης μπορεί να εφαρμοστεί στη θεραπεία όλων των αθλητών. Επί πλέον, αφού είναι λογικό να συμπεράνουμε ότι τα διαστρέμματα του αστραγάλου τα οποία δεν σχετίζονται με αθλήματα είναι συγκρίσιμα με αυτά που σχετίζονται με τα αθλήματα, το πρόγραμμα θα μπορούσε να ωφελήσει το γενικό πληθυσμό».<sup>260</sup>

«Αυτό αναπαράγει και διευρύνει την εργασία των Clarke et al και καταδεικνύει ότι αυτό το CB (cognitive behavioural, γνωστικό συμπεριφορικό) πρόγραμμα πρόληψης μπορεί να

εφαρμοστεί αξιόπιστα και αποτελεσματικά σε διαφορετικές περιπτώσεις από κλινικούς εκτός της ομάδας η οποία αρχικά ανέπτυξε την παρέμβαση. Το μέγεθος της επίδρασης ήταν συνεπές με αυτά που είχαν αναφερθεί παλαιότερα, σε μονο-κεντρικές μελέτες που διερευνούσαν την πρόληψη κατάθλιψης, και ήταν εύρωστο σε όλα τα κέντρα όσον αφορά στις καταθλιπτικές διαταραχές και στα συμπτώματα... Σε αυτή τη μελέτη γενικευσιμότητας, επιλέξαμε μια συγκριτική κατάσταση που είναι σχετική με τη δημόσια υγεία-κοινή φροντίδα... Το δείγμα αποτελείτο κατά κύριο λόγο από άτομα της εργατικής έως της μέσης τάξης, με πρόσβαση σε ασφάλεια υγείας. Δεδομένων των ενδείξεων ότι η CB θεραπεία είναι πιο αποτελεσματική σε εφήβους από σπίτια με μεγαλύτερο εισόδημα, είναι σημαντικό να δοκιμαστούν οι επιδράσεις αυτού του προγράμματος πρόληψης σε ποικίλα οικονομικά και εθνικά δείγματα».<sup>261</sup>

*Εξήγηση:* Η εξωτερική εγκυρότητα ή, διαφορετικά, γενικευσιμότητα ή εφαρμοσιμότητα, είναι ο βαθμός στον οποίο τα αποτελέσματα μιας μελέτης μπορούν να γενικευτούν σε άλλες συνθήκες.<sup>262</sup> Η εσωτερική εγκυρότητα, που αναφέρεται στο βαθμό στον οποίο ο σχεδιασμός και η διενέργεια της δοκιμής εξαλείφουν την πιθανότητα σφαλμάτων, είναι αναγκαία προϋπόθεση για την εξωτερική εγκυρότητα: Τα αποτελέσματα μιας εσφαλμένης δοκιμής είναι άκυρα και το ερώτημα της εξωτερικής εγκυρότητας γίνεται άσχετο. Δεν υπάρχει απόλυτη εξωτερική εγκυρότητα. Ο όρος έχει σημασία μόνο σε σχέση με σαφώς καθοριζόμενες συνθήκες, οι οποίες δεν εξετάστηκαν άμεσα στη δοκιμή. Τα αποτελέσματα μπορούν να γενικευτούν σε έναν επιμέρους συμμετέχοντα ή ομάδες που διαφέρουν από αυτούς που εγγράφηκαν στη δοκιμή σε σχέση με την ηλικία, το φύλο, τη σοβαρότητα της νόσου και τις συνυπάρχουσες καταστάσεις; Είναι τα αποτελέσματα εφαρμόσιμα σε άλλα φάρμακα της ίδιας κατηγορίας, σε άλλη δοσολογία, χρόνο και οδό χορήγησης και μαζί με διαφορετικές ταυτόχρονες θεραπείες; Μπορούμε να αναμένουμε παρόμοια αποτελέσματα στο πρωτοβάθμιο, στο δευτεροβάθμιο και στο τριτοβάθμιο επίπεδο φροντίδας; Τι συμβαίνει με την επίδραση σε σχετικές εκβάσεις οι οποίες δεν εκτιμήθηκαν στη δοκιμή, και με τη σημασία της διάρκειας παρακολούθησης και θεραπείας, ειδικά όσον αφορά στις βλαβερές συνέπειες;<sup>263</sup>

Η εξωτερική εγκυρότητα είναι θέμα κρίσης και εξαρτάται από τους συμμετέχοντες στη δοκιμή, τις συνθήκες της δοκιμής, τα θεραπευτικά σχήματα που ελέγχονται και την έκβαση που εκτιμάται.<sup>5,136</sup> Επομένως, είναι ζωτικής σημασίας να περιγραφούν επαρκείς πληροφορίες σχετικά με τα κριτήρια καταλληλότητας, καθώς και τις συνθήκες και την τοποθεσία (βλέπε στοιχείο 4β), τις παρεμβάσεις και πώς χορηγήθηκαν (βλέπε στοιχείο 5), τον ορισμό των εκβάσεων (βλέπε στοιχείο 6), καθώς και την περίοδο

στρατολόγησης και παρακολούθησης (βλέπε στοιχείο 14). Το ποσοστό των συμμετεχόντων της ομάδας ελέγχου στην οποία εμφανίστηκε η έκβαση (κίνδυνος ομάδας ελέγχου) είναι επίσης σημαντικό. Το ποσοστό των κατάλληλων συμμετεχόντων οι οποίοι αρνήθηκαν να εισαχθούν στη δοκιμή, όπως υποδεικνύεται από το διάγραμμα ροής, σχετίζεται με τη γενικευσιμότητα της δοκιμής, καθώς μπορεί να υποδεικνύει προτιμήσεις για την αποδοχή μιας παρέμβασης. Παρόμοιες θεωρήσεις μπορεί να απευθύνονται και σε προτιμήσεις εκ μέρους των κλινικών ιατρών.<sup>264,265</sup>

Διάφορα θέματα είναι σημαντικά, όταν τα αποτελέσματα μιας δοκιμής εφαρμόζονται σε ένα μεμονωμένο ασθενή.<sup>266–268</sup> Παρ' όλο που πάντα αναμένεται κάποια μεταβολή στην ανταπόκριση στη θεραπεία μεταξύ ενός μεμονωμένου ασθενούς και των ασθενών σε μια δοκιμή ή σε συστηματική ανασκόπηση, οι διαφορές τείνουν να αφορούν στο μέγεθος παρά στην κατεύθυνση.

Παρ' όλο που υπάρχουν σημαντικές εξαιρέσεις,<sup>268</sup> οι θεραπείες (ιδιαίτερα τα φάρμακα<sup>269</sup>) που έχουν βρεθεί ωφέλιμες σε ένα στενό εύρος ασθενών γενικά έχουν ευρύτερη εφαρμογή στην πραγματική πρακτική. Έχουν προταθεί διάφορα πλαίσια για την εκτίμηση της εξωτερικής εγκυρότητας, όπως η ακέραια «διαδικασία αξιολόγησης»<sup>270</sup> και οι κατάλογοι.<sup>271</sup> Μέτρα τα οποία ενσωματώνουν τον αρχικό κίνδυνο στον υπολογισμό του θεραπευτικού αποτελέσματος, όπως ο αριθμός των ατόμων που θα πρέπει να θεραπευτούν για να επιτευχθεί μία επί πλέον ευνοϊκή έκβαση και ο αριθμός των ατόμων που πρέπει να θεραπευτούν για να παρουσιαστεί μία ανεπιθύμητη ενέργεια, είναι επικουρικά στην εκτίμηση της ισορροπίας οφέλους προς τον κίνδυνο (benefit-to-risk) σε ένα μεμονωμένο ασθενή ή σε μια ομάδα με χαρακτηριστικά τα οποία διαφέρουν από τον τυπικό συμμετέχοντα στη μελέτη.<sup>268,272,273</sup> Τελικά, μετά την εξαγωγή ασθενοκεντρικών εκτιμήσεων για πιθανά οφέλη και βλαβερές συνέπειες από μια παρέμβαση, ο κλινικός ιατρός πρέπει να τις ενσωματώσει στις αξίες του ασθενούς και στις προτιμήσεις του για θεραπεία. Παρόμοιες θεωρήσεις ισχύουν κατά την εκτίμηση της γενικευσιμότητας των αποτελεσμάτων σε διαφορετικές συνθήκες και παρεμβάσεις.

### 8.3. Στοιχείο 22. Ερμηνεία συνεπής με τα αποτελέσματα, ισορροπώντας τα οφέλη και τις βλαβερές συνέπειες και λαμβάνοντας υπ' όψη άλλες σχετικές ενδείξεις

*Παράδειγμα:* «Μελέτες οι οποίες δημοσιεύτηκαν πριν από το 1990 πρότειναν ότι η προφυλακτική ανοσοθεραπεία επίσης μειώνει τις νοσοκομειακές λοιμώξεις σε νεογνά πολύ χαμηλού βάρους γέννησης. Εν τούτοις, αυτές οι με-

λέτες ενέγραψαν μικρό αριθμό ασθενών, χρησιμοποίησαν ποικίλους σχεδιασμούς, παρασκευάσματα και δόσεις, και περιελάμβαναν διαφορετικούς πληθυσμούς στις μελέτες. Σε αυτή τη μεγάλη, πολυκεντρική, τυχαιοποιημένη, ελεγχόμενη δοκιμή, η επαναλαμβανόμενη προφυλακτική χορήγηση ενδοφλέβιας ανοσοσφαιρίνης απέτυχε να μειώσει σημαντικά την επίπτωση των νοσοκομειακών λοιμώξεων στα πρόωρα νεογνά με βάρος γέννησης 501–1.500 g». <sup>274</sup>

*Εξήγηση:* Οι αναγνώστες θα επιθυμούν να γνωρίζουν πώς σχετίζονται τα αποτελέσματα της παρούσας μελέτης με εκείνα άλλων ΤΚΔ. Αυτό επιτυγχάνεται καλύτερα, περιλαμβάνοντας μια τυπική συστηματική ανασκόπηση στο τμήμα των αποτελεσμάτων ή της συζήτησης της αναφοράς. <sup>83,275–277</sup> Η εν λόγω σύνθεση μπορεί να μην είναι πρακτική για τους συγγραφείς της μελέτης, αλλά είναι συχνά εφικτή η παράθεση μιας συστηματικής ανασκόπησης παρόμοιων δοκιμών. Η συστηματική ανασκόπηση θα βοηθήσει τους αναγνώστες να εκτιμήσουν εάν τα αποτελέσματα της ΤΚΔ είναι παρόμοια με εκείνα άλλων δοκιμών στο ίδιο γνωστικό αντικείμενο και εάν οι συμμετέχοντες ήταν παρόμοιοι σε άλλες μελέτες. Οι δημοσιεύσεις των ΤΚΔ συχνά δεν αντιμετωπίζουν επαρκώς τα σχετικά σημεία. Μπεύζιανές μέθοδοι μπορούν να χρησιμοποιηθούν για το στατιστικό συνδυασμό των δεδομένων της δοκιμής με προηγούμενες ενδείξεις. <sup>278</sup>

Συστήνουμε τουλάχιστον η συζήτηση να είναι όσο το δυνατόν πιο συστηματική και να βασίζεται σε περιεκτική έρευνα, παρά να περιορίζεται σε μελέτες οι οποίες υποστηρίζουν τα αποτελέσματα της παρούσας μελέτης. <sup>279</sup>

## 9. ΑΛΛΕΣ ΠΛΗΡΟΦΟΡΙΕΣ

### 9.1. Στοιχείο 23. Αριθμός καταχώρησης και όνομα του μητρώου δοκιμών

*Παράδειγμα:* «Η μελέτη καταχωρήθηκε στο ClinicalTrials.gov, με αριθμό NCT00244842». <sup>280</sup>

*Εξήγηση:* Οι συνέπειες της μη δημοσίευσης ολόκληρων των δοκιμών, <sup>281,282</sup> της επιλεκτικής αναφοράς των εκβάσεων στο πλαίσιο της δοκιμής και της ανάλυσης σύμφωνα με το πρωτόκολλο, παρά της ανάλυσης πρόθεση για θεραπεία, έχουν τεκμηριωθεί καλά. <sup>55,56,283</sup> Η συγκαλυμμένη περιττή δημοσίευση των κλινικών δοκιμών μπορεί επίσης να δημιουργήσει προβλήματα, ειδικά για τους συγγραφείς των συστηματικών ανασκοπήσεων, όταν αποτελέσματα από την ίδια δοκιμή, από απροσεξία, περιλαμβάνονται περισσότερες από μία φορές. <sup>284</sup>

Για να ελαχιστοποιηθούν ή να αποφευχθούν τα προβλήματα αυτά, έχουν γίνει επαναλαμβανόμενες εκκλήσεις

τα τελευταία 25 χρόνια για καταχώρηση των κλινικών δοκιμών κατά την έναρξή τους, για ανάθεση μοναδικών αριθμών αναγνώρισης της μελέτης και για καταγραφή άλλων βασικών πληροφοριών σχετικά με τη μελέτη, έτσι ώστε ουσιώδεις πληροφορίες να καταστούν δημόσια προσβάσιμες. <sup>285–288</sup> Τα πρόσφατα σοβαρά προβλήματα παρακράτησης δεδομένων <sup>289</sup> οδήγησαν σε μια ανανεωμένη προσπάθεια καταχώρησης των κλινικών δοκιμών. Πράγματι, ο Παγκόσμιος Οργανισμός Υγείας δηλώνει ότι «η καταχώρηση όλων των παρεμβατικών δοκιμών είναι μια επιστημονική, δεοντολογική και ηθική υπευθυνότητα» ([www.who.int/ictpr/en](http://www.who.int/ictpr/en)). Με την καταχώρηση μιας τυχαιοποιημένης δοκιμής, οι συγγραφείς τυπικά αναφέρουν μια ελάχιστη σειρά πληροφοριών, ενώ αποκτάται ένας μοναδικός αριθμός καταχώρησης της δοκιμής.

Το Σεπτέμβριο του 2004, η Διεθνής Επιτροπή των Εκδοτών Ιατρικών Περιοδικών (International Committee of Medical Journal Editors, ICMJE) άλλαξε την πολιτική της, αναφέροντας ότι θα θεωρούνται προς δημοσίευση δοκιμές μόνο εάν είχαν καταχωρηθεί πριν από την εγγραφή του πρώτου συμμετέχοντα. <sup>290</sup> Αυτό οδήγησε σε μια δραματική αύξηση του αριθμού των δοκιμών που καταχωρήθηκαν. <sup>291</sup> Η ICMJE παρέχει καθοδήγηση για τα αποδεκτά μητρώα ([www.icmje.org/faq.pdf](http://www.icmje.org/faq.pdf)).

Σε μια πρόσφατη έρευνα 165 οδηγιών προς τους συγγραφείς από ιατρικά περιοδικά με υψηλό συντελεστή απήχησης, τα 44 περιοδικά δήλωναν σαφώς, ως απαίτηση για δημοσίευση στο περιοδικό, όλες οι πρόσφατες κλινικές δοκιμές να είναι καταχωρημένες. <sup>292</sup>

Οι συγγραφείς πρέπει να παρέχουν το όνομα του μητρώου και το μοναδικό αριθμό καταχώρησης. Αν οι συγγραφείς δεν καταχώρησαν τη δοκιμή τους, πρέπει να το αναφέρουν σαφώς, επεξηγώντας το λόγο.

### 9.2. Στοιχείο 24. Πού μπορεί να είναι προσβάσιμο το συνολικό πρωτόκολλο της δοκιμής, εάν είναι διαθέσιμο

*Παράδειγμα:* «Πλήρεις λεπτομέρειες σχετικά με το πρωτόκολλο της δοκιμής μπορούν να ανευρεθούν στο Συμπληρωματικό Παράρτημα, το οποίο είναι διαθέσιμο με το πλήρες κείμενο του άρθρου στο [www.nejm.org](http://www.nejm.org)». <sup>293</sup>

*Εξήγηση:* Ένα πρωτόκολλο ολόκληρης της μελέτης (παρά ένα πρωτόκολλο μιας συγκεκριμένης διαδικασίας μέσα στη δοκιμή) είναι σημαντικό, γιατί προκαθορίζει τις μεθόδους της τυχαιοποιημένης δοκιμής, όπως την κύρια έκβαση (βλέπε στοιχείο 6α). Η ύπαρξη ενός πρωτοκόλλου μπορεί να βοηθήσει στον περιορισμό της πιθανότητας των εκ των υστέρων αδήλων αλλαγών στις μεθόδους της



δοκιμής και στην επιλεκτική αναφορά έκβασης (βλέπε στοιχείο ββ). Τα στοιχεία που μπορεί να είναι σημαντικά προκειμένου να συμπεριληφθούν στο πρωτόκολλο μιας τυχαίοποιημένης δοκιμής περιγράφονται αλλού.<sup>294</sup>

Υπάρχουν πολλές επιλογές για τους συγγραφείς, αν πρέπει να αναλογιστούν τη διασφάλιση του να καταστήσουν το πρωτόκολλο της δοκιμής διαθέσιμο για ενδιαφερόμενους αναγνώστες. Όπως περιγράφεται στο παραπάνω παράδειγμα, τα περιοδικά που αναφέρουν τα κύρια αποτελέσματα μιας δοκιμής μπορούν να καταστήσουν το πρωτόκολλο της δοκιμής διαθέσιμο στην ιστοσελίδα τους. Η διαθεσιμότητα των αποτελεσμάτων της δοκιμής και του πρωτοκόλλου ενισχύεται, όταν το περιοδικό έχει ελεύθερη πρόσβαση. Μερικά περιοδικά (όπως το *Trials*) δημοσιεύουν τα πρωτόκολλα των δοκιμών και σε μια τέτοια δημοσίευση μπορεί να γίνει μνεία, όταν αναφέρονται τα κύρια αποτελέσματα της δοκιμής.

Η καταχώρηση της δοκιμής (βλέπε στοιχείο 23) μπορεί ακόμη να διασφαλίσει ότι πολλές λεπτομέρειες σχετικά με το πρωτόκολλο της δοκιμής είναι διαθέσιμες, εφ' όσον τα ελάχιστα χαρακτηριστικά της δοκιμής που περιλαμβάνονται σε μια εγκεκριμένη βάση καταχώρησης δοκιμών περιέχουν αρκετά στοιχεία του πρωτοκόλλου και αποτελέσματα ([www.who.int/ictpr/en](http://www.who.int/ictpr/en)). Οι ερευνητές των δοκιμών μπορούν να τοποθετήσουν το πρωτόκολλο της δοκιμής σε μια ιστοσελίδα, μέσω του εργοδότη τους. Όποιος μηχανισμός και αν χρησιμοποιηθεί, ενθαρρύνουμε όλους τους ερευνητές των δοκιμών να καταστήσουν το πρωτόκολλό τους εύκολα προσβάσιμο σε όλους τους ενδιαφερόμενους αναγνώστες.

### 9.3. Στοιχείο 25. Πηγές χρηματοδότησης και άλλη υποστήριξη (όπως η προμήθεια φαρμάκων), ρόλος των χρηματοδοτών

*Παράδειγμα:* «Υποστήριξη από χορηγία δόθηκε για την παρέμβαση από το Plan International και για την έρευνα από το Wellcome Trust και το Κοινό Πρόγραμμα των Ηνωμένων Εθνών σχετικά με το HIV/AIDS (UNAIDS). Οι χορηγοί δεν είχαν κάποιο ρόλο στο σχεδιασμό της μελέτης, στη συλλογή των δεδομένων, στην απόφαση για δημοσίευση ή στην προετοιμασία του χειρογράφου».<sup>295</sup>

«Η μελέτη αυτή χρηματοδοτήθηκε από τη φαρμακευτική εταιρεία GlaxoSmithKline. Η GlaxoSmithKline ενεπλάκη στο σχεδιασμό και στη διεξαγωγή της μελέτης και παρείχε λογιστική υποστήριξη κατά τη διάρκεια της δοκιμής. Οι υπάλληλοι του χορηγού εργάστηκαν μαζί με τους ερευνητές στην προετοιμασία του σχεδίου στατιστικής ανάλυσης, αλλά η ανάλυση πραγματοποιήθηκε στο Πανεπιστήμιο της Utah.

Η εργασία ετοιμάστηκε από τον Dr Shaddy και τα μέλη της οργανωτικής επιτροπής. Η GlaxoSmithKline είχε την άδεια να κρίνει την εργασία και να προτείνει μεταβολές. Όμως, η τελική απόφαση σχετικά με το περιεχόμενο αφορούσε αποκλειστικά στους συγγραφείς».<sup>296</sup>

*Εξήγηση:* Οι συγγραφείς πρέπει να αναφέρουν τις πηγές χρηματοδότησης της δοκιμής, καθ' όσον αυτό αποτελεί σημαντική πληροφορία για τους αναγνώστες οι οποίοι εκτιμούν μια δοκιμή. Μελέτες έχουν δείξει ότι η έρευνα που χρηματοδοτείται από τη φαρμακευτική βιομηχανία είναι πιο πιθανό να παράγει αποτελέσματα που ευνοούν το προϊόν το οποίο κατασκευάζει η εταιρεία που χρηματοδοτεί την έρευνα, σε σύγκριση με άλλες μελέτες οι οποίες χρηματοδοτούνται από άλλες πηγές.<sup>279–300</sup> Σε μια συστηματική ανασκόπηση 30 μελετών σχετικά με τη χρηματοδότηση βρέθηκε ότι η έρευνα που χρηματοδοτείται από τη φαρμακευτική βιομηχανία είχε τέσσερις φορές μεγαλύτερη πιθανότητα να έχει αποτελέσματα σε όφελος του χορηγού έναντι της έρευνας που χρηματοδοτείται από άλλες πηγές (λόγος αναλογιών, odds ratio 4,05, 95% διαστήματα εμπιστοσύνης 2,98–5,51).<sup>297</sup> Μια μεγάλη αναλογία των δημοσιεύσεων των δοκιμών αυτή τη στιγμή δεν αναφέρει πηγές χρηματοδότησης. Ο βαθμός της μειωμένης αναφοράς είναι δύσκολο να ποσοτικοποιηθεί. Μια έρευνα 370 φαρμακευτικών δοκιμών βρήκε ότι στο 29% υπήρξε αποτυχία στην αναφορά των πηγών χρηματοδότησης.<sup>301</sup> Σε μια άλλη έρευνα των καταχωρημένων τυχαίοποιημένων δοκιμών του PubMed που δημοσιεύτηκαν το Δεκέμβριο του 2000, η πηγή χρηματοδότησης αναφερόταν στο 66% από τις 519 δοκιμές.<sup>16</sup>

Το επίπεδο της εμπλοκής από πλευράς του χρηματοδότη και η επίδρασή του στο σχεδιασμό, στη διενέργεια, στην ανάλυση και στην αναφορά των δοκιμών ποικίλλει. Γι' αυτό, είναι σημαντικό οι συγγραφείς να περιγράφουν με λεπτομέρειες το ρόλο των χρηματοδοτών. Αν ο χρηματοδότης δεν έχει τέτοιου είδους εμπλοκή, οι συγγραφείς θα πρέπει να το αναφέρουν. Παρομοίως, οι συγγραφείς πρέπει να αναφέρουν άλλες πηγές υποστήριξης, όπως η προμήθεια και η προετοιμασία των φαρμάκων ή του εξοπλισμού, ή στην ανάλυση των δεδομένων και στη συγγραφή του χειρογράφου.<sup>302</sup>

## 10. ΑΝΑΦΕΡΟΝΤΑΣ ΤΙΣ ΤΥΧΑΙΟΠΟΙΗΜΕΝΕΣ ΚΛΙΝΙΚΕΣ ΔΟΚΙΜΕΣ ΟΙ ΟΠΟΙΕΣ ΔΕΝ ΕΙΧΑΝ ΣΧΕΔΙΑΣΜΟ ΔΥΟ ΠΑΡΑΛΛΗΛΩΝ ΟΜΑΔΩΝ

Οι συστάσεις του CONSORT εστιάζονται πρωταρχικά στις ΤΚΔ, με παράλληλο σχεδιασμό και δύο ομάδες θεραπείας. Οι περισσότερες ΤΚΔ έχουν αυτόν το σχεδιασμό, και μια ουσιαστική μειονότητα δεν τον διαθέτουν: το 45%

(233/519) των ΤΚΔ οι οποίες δημοσιεύτηκαν το Δεκέμβριο του 2000 και το 39% (242/616) αυτών που δημοσιεύτηκαν το Δεκέμβριο του 2006.

Το μεγαλύτερο τμήμα της δήλωσης CONSORT έχει εφαρμογή ισότιμα σε όλους τους σχεδιασμούς των μελετών, αλλά υπάρχουν μερικά επί πλέον θέματα που αφορούν σε κάθε σχεδιασμό. Πριν από τη δημοσίευση της αναθεωρημένης δήλωσης CONSORT 2001, η ομάδα CONSORT αποφάσισε να αναπτύξει επεκτάσεις της κύριας δήλωσης CONSORT, σχετικές με ειδικούς σχεδιασμούς δοκιμών. Επεκτάσεις έχουν δημοσιευτεί σε σχέση με την αναφορά των τυχαιοποιημένων δοκιμών συρροής, αλλά και των δοκιμών μη υπεροχής και ισοδυναμίας. Η έλλειψη πόρων ήταν η αιτία που δεν ολοκληρώθηκαν άλλες προγραμματισμένες επεκτάσεις: θα καλύψουν δοκιμές με τους ακόλουθους σχεδιασμούς: Παράλληλο πολλαπλών σκελών (multiarm), παραγοντικό, διασταυρούμενο, εντός ατόμου (within-person).

Οι συγγραφείς οι οποίοι αναφέρουν μια δοκιμή με σχεδιασμό συρροής ή χρησιμοποιούν ένα πλαίσιο εργασίας μη υπεροχής ή ισοδυναμίας πρέπει να συμβουλευονται επεκτάσεις του CONSORT, συμπληρωματικά εκείνων του συγκεκριμένου εγγράφου. Σε αυτό το σημείο παρατίθενται μερικά ενδιάμεσα σχόλια σχετικά με τους άλλους σχεδιασμούς. Σε κάθε περίπτωση, ο σχεδιασμός της δοκιμής πρέπει να καθίσταται σαφής τόσο στο κυρίως κείμενο, όσο και στην περίληψη του άρθρου.

Οι δοκιμές παράλληλων πολλαπλών σκελών (>2 ομάδες) χρειάζονται τη μικρότερη τροποποίηση των τυπικών οδηγιών του CONSORT. Το διάγραμμα ροής εύκολα μπορεί να επεκταθεί. Η κύρια διαφορά από δοκιμές με δύο ομάδες σχετίζεται με τη διασαφήνιση του πώς η υπόθεση της μελέτης έχει σχέση με τις πολλαπλές ομάδες και τις συνεπακόλουθες μεθόδους ανάλυσης των δεδομένων και της ερμηνείας. Για παραγοντικές μελέτες, η πιθανότητα αλληλεπίδρασης μεταξύ των παρεμβάσεων πρέπει γενικά να ληφθεί υπ' όψη. Επί πλέον των συνολικών συγκρίσεων των συμμετεχόντων, οι οποίοι έλαβαν ή δεν έλαβαν κάθε παρέμβαση υπό μελέτη, οι ερευνητές πρέπει να θεωρήσουν ακόμη την αναφορά των αποτελεσμάτων για κάθε θεραπευτικό συνδυασμό.<sup>303</sup>

Στις διασταυρούμενες δοκιμές, κάθε συμμετέχων λαμβάνει δύο (ή περισσότερες θεραπείες) με τυχαία σειρά. Το κύριο επί πλέον ζήτημα για θεώρηση σχετίζεται με την εξαρτημένη φύση των δεδομένων, η οποία επηρεάζει το σχεδιασμό και την ανάλυση.<sup>304</sup> Παρόμοια ζητήματα επηρεάζουν τις συγκρίσεις εντός ατόμου, στις οποίες οι συμμετέχοντες λαμβάνουν δύο θεραπείες ταυτόχρονα (συνήα, σε αντιστοιχισμένα όργανα). Επίσης, εξ αιτίας του κινδύνου

χρονικών και συστηματικών υπολειπόμενων επιδράσεων (carryover effects), αντίστοιχα, και στις δύο περιπτώσεις η επιλογή του σχεδιασμού χρειάζεται δικαιολόγηση.

Η ομάδα CONSORT σκοπεύει να δημοσιεύσει επεκτάσεις του CONSORT για να καλύψει όλους αυτούς τους σχεδιασμούς. Επί πλέον, θα δημοσιεύσουμε αναθεωρήσεις στις υπάρχουσες οδηγίες για τυχαιοποιημένες δοκιμές συρροής και δοκιμές μη υπεροχής και ισοδυναμίας, ώστε να λάβουμε υπ' όψη αυτή τη μεγάλη αναθεώρηση των γενόσημων οδηγιών του CONSORT.

## 11. ΣΥΖΗΤΗΣΗ

Η εκτίμηση των παρεμβάσεων στη φροντίδα υγείας μπορεί να είναι παραπλανητική, εκτός εάν οι ερευνητές διασφαλίζουν αμερόληπτες συγκρίσεις. Η τυχαία διανομή σε ομάδες μελέτης παραμένει η μόνη μέθοδος η οποία ελαχιστοποιεί σφάλματα λόγω συγχυτικών παραγόντων και σφάλματα επιλογής. Μη τυχαιοποιημένες δοκιμές τείνουν να καταλήγουν σε εκτίμηση μεγαλύτερων επιδράσεων της θεραπείας από τις τυχαιοποιημένες δοκιμές.<sup>305,306</sup>

Εν τούτοις, τα σφάλματα θέτουν σε κίνδυνο ακόμη και τις ΤΚΔ, εάν οι ερευνητές διενεργούν τέτοιες δοκιμές με ακατάλληλο τρόπο.<sup>307</sup> Μια πρόσφατη συστηματική ανασκόπηση, η οποία συναθροίζει τα αποτελέσματα διαφόρων μεθοδολογικών ερευνών, εντόπισε ότι για υποκειμενικές εκβάσεις οι δοκιμές που χρησιμοποιούσαν ανεπαρκή ή ασαφή απόκρυψη διανομής απέφεραν 31% μεγαλύτερη εκτίμηση της επίδρασης από τις δοκιμές με επαρκή απόκρυψη, ενώ οι δοκιμές που δεν ήταν τυφλές απέφεραν 25% μεγαλύτερες εκτιμήσεις.<sup>153</sup> Όπως θα μπορούσε να αναμένεται, υπήρχε ισχυρή σχέση ανάμεσα στις δύο συνθήκες.

Ο σχεδιασμός και η εκτέλεση μιας ΤΚΔ απαιτεί τόσο μεθοδολογική, όσο και κλινική εμπειρογνωμοσύνη, σχολαστική προσπάθεια<sup>143,308</sup> και ένα υψηλό επίπεδο ετοιμότητας για απρόβλεπτες δυσκολίες. Οι δημοσιεύσεις των ΤΚΔ θα πρέπει να γράφονται με ανάλογα μεγάλη προσοχή προς μείωση των σφαλμάτων. Οι αναγνώστες δεν θα πρέπει να εικάζουν: οι μέθοδοι πρέπει να είναι ολοκληρωμένες και διαφανείς, έτσι ώστε οι αναγνώστες να μπορούν εύκολα να ξεχωρίσουν τις δοκιμές με τα αμερόληπτα αποτελέσματα από εκείνες με τα αμφισβητήσιμα αποτελέσματα. Η έγκυρη επιστήμη περιλαμβάνει την επαρκή αναφορά και η διενέργεια των δεοντολογικών δοκιμών στηρίζεται στη βάση της έγκυρης επιστήμης.<sup>309</sup>

Ελπίζουμε ότι αυτή η αναθεώρηση του επεξηγηματικού άρθρου του CONSORT θα συνδράμει τους συγγραφείς στη χρήση της έκδοσης του 2010 του CONSORT και θα εξηγήσει

με γενικούς όρους τη σημασία της επαρκούς αναφοράς στις δοκιμές. Η δήλωση CONSORT μπορεί να βοηθήσει τους ερευνητές να σχεδιάσουν δοκιμές στο μέλλον<sup>310</sup> και να καθοδηγήσει τους κριτές και τους εκδότες στην εκτίμηση των χειρογράφων. Ιδιαίτερα ενθαρρύνουμε τους κριτές και τους εκδότες να χρησιμοποιούν τον κατάλογο του CONSORT για να εκτιμούν αν οι συγγραφείς ανέφεραν τα αποτελέσματά τους σύμφωνα με τα σχετικά στοιχεία. Τέτοιες εκτιμήσεις είναι πιθανό να βελτιώσουν τη σαφήνεια και τη διαφάνεια των δημοσιευμένων δοκιμών. Επειδή το CONSORT είναι ένα εξελισσόμενο έγγραφο, απαιτείται μια δυναμική διαδικασία συνεχούς εκτίμησης, εκκαθάρισης και, εάν κρίνεται αναγκαίο, αλλαγής. Έτσι, λοιπόν, χρειάστηκε η παρούσα αναθεώρηση του καταλόγου και του επεξηγηματικού άρθρου. Όσο θα συσσωρεύονται νέες ενδείξεις και κριτικά σχόλια, θα αξιολογείται η ανάγκη για μελλοντικές αναθεωρήσεις.

Η πρώτη εκδοχή της δήλωσης CONSORT το 1996 φαίνεται να έχει οδηγήσει σε βελτίωση της ποιότητας των αναφορών των ΤΚΔ στα περιοδικά που το υιοθέτησαν.<sup>50–54</sup> Άλλες ομάδες χρησιμοποιούν το πρότυπο του CONSORT για τη βελτίωση των αναφορών και των άλλων ερευνητικών σχεδιασμών, όπως οι διαγνωστικές δοκιμασίες<sup>311</sup> και οι περιγραφικές μελέτες.<sup>312</sup>

Η ιστοσελίδα του CONSORT ([www.consort-statement.org](http://www.consort-statement.org)) ιδρύθηκε για να παρέχει εκπαιδευτικό υλικό και μια αποθηκευτική βάση δεδομένων υλικού σχετικού με τις ΤΚΔ. Η ιστοσελίδα περιλαμβάνει πολλά παραδείγματα από πραγματικές δοκιμές, περιλαμβανομένων όλων των παραδειγμάτων τα οποία περιέχονται στο παρόν άρθρο. Θα συνεχιστεί η προσθήκη καλών και κακών παραδειγμάτων αναφορών στη βάση δεδομένων, και καλούνται οι αναγνώστες να υποβάλλουν περαιτέρω προτάσεις, ερχόμενοι σε επικοινωνία μαζί μας μέσω της ιστοσελίδας. Η ομάδα του CONSORT θα συνεχίσει να εξετάζει τη βιβλιογραφία, ώστε να βρει σχετικά άρθρα τα οποία ασχολούνται με την αναφορά των ΤΚΔ, ενώ οι συγγραφείς τέτοιων άρθρων καλούνται να μας ενημερώνουν σχετικά. Όλες αυτές οι πληροφορίες θα είναι προσβάσιμες μέσω της ιστοσελίδας του CONSORT, η οποία ενημερώνεται τακτικά.

Περισσότερα από 400 κορυφαία γενικά ιατρικά περιοδικά, περιοδικά ειδικότητων και βιοϊατρικές εκδοτικές ομάδες, περιλαμβανομένου του ICMJE, της Παγκόσμιας Εταιρείας των Εκδοτών Ιατρικών Περιοδικών και του Συμβουλίου των Εκδοτών Επιστήμης, έχουν δώσει την επίσημη υποστήριξή τους στο CONSORT. Καλούμε και άλλα περιοδικά, τα οποία τα ενδιαφέρει η ποιότητα της αναφοράς των κλινικών δοκιμών, να υποστηρίξουν τη δήλωση CONSORT και να επικοινωνήσουν μαζί μας μέσω

της ιστοσελίδας μας για να μας καταστήσουν γνωστή την υποστήριξή τους. Οι απώτεροι ευεργετημένοι από αυτή την ομαδική προσπάθεια πρέπει να είναι οι άνθρωποι, οι οποίοι για οποιονδήποτε λόγο χρειάζονται μια παρέμβαση από την κοινότητα της φροντίδας υγείας.

## ΕΥΧΑΡΙΣΤΙΕΣ

*Είμαστε ευγνώμονες στον Frank Davidoff και στον Tom Lang για την εμπλοκή τους στην έκδοση του 2001, του εγγράφου επεξήγησης και επεξεργασίας του CONSORT. Ιδιαίτερες ευχαριστίες στη Mary Ocampro, συντονίστρια του Ottawa CONSORT, η οποία συνέδραμε στη διανομή αυτού του εγγράφου επεξήγησης και επεξεργασίας, καθώς και της δήλωσης CONSORT.*

## Οι συμβάλλοντες της ομάδας CONSORT για το CONSORT 2010

Douglas G. Altman, Centre for Statistics in Medicine, University of Oxford, Ηνωμένο Βασίλειο, Virginia Barbour, PLoS Medicine, Ηνωμένο Βασίλειο, Jesse A. Berlin, Johnson & Johnson Pharmaceutical Research and Development, ΗΠΑ, Isabelle Boutron, University Paris 7 Denis Diderot, Assistance Publique des Hôpitaux de Paris, INSERM, Γαλλία, P.J. Devereaux, McMaster University, Καναδάς, Kay Dickersin, Johns Hopkins Bloomberg School of Public Health, ΗΠΑ, Diana Elbourne, London School of Hygiene & Tropical Medicine, Ηνωμένο Βασίλειο, Susan Ellenberg, University of Pennsylvania School of Medicine, ΗΠΑ, Val Gebski, University of Sydney, Αυστραλία, Steven Goodman, Journal of the Society for Clinical Trials, ΗΠΑ, Peter C. Gøtzsche, Nordic Cochrane Centre, Δανία, Trish Groves, BMJ, Ηνωμένο Βασίλειο, Steven Grunberg, American Society of Clinical Oncology, ΗΠΑ, Brian Haynes, McMaster University, Καναδάς, Sally Hopewell, Centre for Statistics in Medicine, University of Oxford, Ηνωμένο Βασίλειο, Astrid James, Lancet, Peter Juhn, Johnson & Johnson, ΗΠΑ, Philippa Middleton, University of Adelaide, Αυστραλία, Don Minckler, University of California Irvine, ΗΠΑ, David Moher, Ottawa Methods Centre, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Καναδάς, Victor M. Montori, Knowledge and Encounter Research Unit, Mayo Clinic College of Medicine, ΗΠΑ, Cynthia Mulrow, Annals of Internal Medicine, ΗΠΑ, Stuart Pocock, London School of Hygiene & Tropical Medicine, Ηνωμένο Βασίλειο, Drummond Rennie, JAMA, ΗΠΑ, David L. Schriger, Annals of Emergency Medicine, ΗΠΑ, Kenneth F. Schulz, Family Health International, ΗΠΑ, Iveta Simera, EQUATOR Network, Ηνωμένο Βασίλειο, Elizabeth Wager, Sideview, Ηνωμένο Βασίλειο.

## Χρηματοδότηση

Με ευγνωμοσύνη αναγνωρίζουμε τη χρηματική υποστήριξη από το Εθνικό Ινστιτούτο για την έρευνα σε θέματα υγείας του Ηνωμένου Βασιλείου, τα Καναδικά Ινστιτούτα για την έρευνα σε θέματα υγείας, την Προεδρική Χρηματοδότηση, την εταιρεία Johnson & Johnson, το BMJ και την Αμερικανική Εταιρεία Κλινικής Ογκολογίας.

Στοχεύοντας στην ενθάρρυνση της διάδοσης της δήλωσης CONSORT 2010, το παρόν άρθρο είναι ελεύθερα προσβάσιμο στο [bmj.com](http://bmj.com) και θα δημοσιευτεί, επίσης, στο *Journal of Clinical Epidemiology*. Το σύνολο των συγγραφέων κατέχουν τα δικαιώματα αυτού του άρθρου. Για λεπτομέρειες σχετικά με μελλοντική χρήση, ανατρέξτε στην ιστοσελίδα του CONSORT ([www.consort-statement.org](http://www.consort-statement.org)).

## ABSTRACT

### CONSORT 2010, explanation and elaboration: Updated guidelines for reporting parallel group randomised trials

GREEK TRANSLATION – EDITION: S. KOKKALI, C. BIRTSOU, A.B. HAIDICH

*Department of Hygiene and Epidemiology, Aristotle University of Thessaloniki, School of Medicine, Thessaloniki, Greece*

*Archives of Hellenic Medicine 2011, 28(Suppl 1):10–59*

Overwhelming evidence shows that the quality of reporting of randomized controlled trials (RCTs) is not optimal. Without transparent reporting, readers can neither judge the reliability and validity of trial findings nor extract valid information for systematic reviews. Recent methodological analyses indicate that inadequate reporting and design are associated with biased estimates of treatment effects. Such systematic error is seriously damaging to RCTs, which are considered the gold standard for evaluating interventions because of their ability to minimize or avoid bias. A group of scientists and editors developed the Consolidated Standards of Reporting Trials (CONSORT) statement to improve the quality of reporting of RCTs. This was first published in 1996 and updated in 2001. The statement consists of a checklist and flow diagram that authors can use for reporting an RCT. Many leading medical journals and major international editorial groups have endorsed the CONSORT statement. The statement facilitates critical appraisal and interpretation of RCTs. During the 2001 CONSORT revision, it became clear that explanation and elaboration of the principles underlying the CONSORT statement would help investigators and others to write or appraise trial reports. A CONSORT explanation and elaboration article was published in 2001 alongside the 2001 version of the CONSORT statement. After an expert meeting in January 2007, the CONSORT statement was further revised and published as the CONSORT 2010 Statement. This update improves the wording and clarity of the previous checklist and incorporates recommendations related to topics that have only recently received recognition, such as selective outcome reporting bias. This explanatory and elaboration document –intended to enhance the use, understanding, and dissemination of the CONSORT statement– has also been extensively revised. It presents the meaning and rationale for each new and updated checklist item, providing examples of good reporting and, where possible, references to relevant empirical studies. Several examples of flow diagrams are included. The CONSORT 2010 statement, this revised explanatory and elaboration document, and the associated website ([www.consort-statement.org](http://www.consort-statement.org)) should be helpful resources for improving the reporting of randomized trials.

**Key words:** Bias, CONSORT statement, Critical appraisal, Quality, Randomized controlled trial (RCT), Systematic error, Transparency

## Βιβλιογραφία

1. RENNIE D. CONSORT revised – improving the reporting of randomized trials. *JAMA* 2001, 285:2006–2007
2. SCHULZ KF, CHALMERS I, HAYES RJ, ALTMAN DG. Empirical evidence of bias. Dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995, 273:408–412
3. MOHER D. CONSORT: An evolving tool to help improve the quality of reports of randomized controlled trials. Consolidated Standards of Reporting Trials. *JAMA* 1998, 279:1489–1491
4. KJAERGARD LL, VILLUMSEN J, GLUUD C. Quality of randomised clinical trials affects estimates of intervention efficacy. 7th Cochrane Colloquium, Rome, Italy, 1999



5. JÜNI P, ALTMAN DG, EGGER M. Systematic reviews in health care: Assessing the quality of controlled clinical trials. *Br Med J* 2001, 323:42–46
6. VELDTHUYZEN VAN ZANTEN SJ, CLEARY C, TALLEY NJ, PETERSON TC, NYRÉN O, BRADLEY LA ET AL. Drug treatment of functional dyspepsia: A systematic analysis of trial methodology with recommendations for design of future trials. *Am J Gastroenterol* 1996, 91:660–673
7. TALLEY NJ, OWEN BK, BOYCE P, PATERSON K. Psychological treatments for irritable bowel syndrome: A critique of controlled treatment trials. *Am J Gastroenterol* 1996, 91:277–283
8. ADETUGBO K, WILLIAMS H. How well are randomized controlled trials reported in the dermatology literature? *Arch Dermatol* 2000, 136:381–385
9. KJAERGARD LL, NIKOLOVA D, GLUUD C. Randomized clinical trials in hepatology: Predictors of quality. *Hepatology* 1999, 30:1134–1138
10. SCHOR S, KARTEN I. Statistical evaluation of medical journal manuscripts. *JAMA* 1966, 195:1123–1128
11. GORE SM, JONES IG, RYTTER EC. Misuse of statistical methods: Critical assessment of articles in BMJ from January to March 1976. *Br Med J* 1977, 1:85–87
12. HALL JC, HILL D, WATTS JM. Misuse of statistical methods in the Australasian surgical literature. *Aust NZ J Surg* 1982, 52:541–543
13. ALTMAN DG. Statistics in medical journals. *Stat Med* 1982, 1:59–71
14. POCKOCK SJ, HUGHES MD, LEE RJ. Statistical problems in the reporting of clinical trials. A survey of three medical journals. *N Engl J Med* 1987, 317:426–432
15. ALTMAN DG. The scandal of poor medical research. *Br Med J* 1994, 308:283–284
16. CHAN AW, ALTMAN DG. Epidemiology and reporting of randomised trials published in PubMed journals. *Lancet* 2005, 365:1159–1162
17. HOPEWELL S, DUTTON S, YU LM, CHAN AW, ALTMAN DG. The quality of reports of randomised trials in 2000 and 2006: Comparative study of articles indexed in PubMed. *Br Med J* 2010, 340:c723
18. HOLLIS S, CAMPBELL F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. *Br Med J* 1999, 319:670–674
19. LAI TY, WONG VW, LAM RF, CHENG AC, LAM DS, LEUNG GM. Quality of reporting of key methodological items of randomized controlled trials in clinical ophthalmic journals. *Ophthalmic Epidemiol* 2007, 14:390–398
20. MOHER D, FORTIN P, JADAD AR, JÜNI P, KLASSEN T, LE LORIER J ET AL. Completeness of reporting of trials published in languages other than English: Implications for conduct and reporting of systematic reviews. *Lancet* 1996, 347:363–366
21. JUNKER CA. Adherence to published standards of reporting: A comparison of placebo-controlled trials published in English or German. *JAMA* 1998, 280:247–249
22. ALTMAN DG. Randomisation. *Br Med J* 1991, 302:1481–1482
23. SCHULZ KF, CHALMERS I, GRIMES DA, ALTMAN DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. *JAMA* 1994, 272:125–128
24. ANONYMOUS. Streptomycin treatment of pulmonary tuberculosis. *Br Med J* 1948, 2:769–782
25. SCHULZ KF. Randomized controlled trials. *Clin Obstet Gynecol* 1998, 41:245–256
26. GREENLAND S. Randomization, statistics, and causal inference. *Epidemiology* 1990, 1:421–429
27. ARMITAGE P. The role of randomization in clinical trials. *Stat Med* 1982, 1:345–352
28. KLEIJNEN J, GØTZSCHE PC, KUNZ R, OXMAN AD, CHALMERS I. So what's so special about randomisation. In: Maynard A, Chalmers I (eds) *Non-random reflections on health services research*. BMJ Books, London, 1997:93–106
29. CHALMERS I. Assembling comparison groups to assess the effects of health care. *J R Soc Med* 1997, 90:379–386
30. NICOLUCCI A, GRILLI R, ALEXANIAN AA, APOLONE G, TORRIV, LIBERATI A. Quality, evolution, and clinical implications of randomized, controlled trials on the treatment of lung cancer. A lost opportunity for meta-analysis. *JAMA* 1989, 262:2101–2107
31. AH-SEE KW, MOLONY NC. A qualitative assessment of randomized controlled trials in otolaryngology. *J Laryngol Otol* 1998, 112:460–463
32. ALTMAN DG, DORÉ CJ. Randomisation and baseline comparisons in clinical trials. *Lancet* 1990, 335:149–153
33. THORNLEY B, ADAMS C. Content and quality of 2000 controlled trials in schizophrenia over 50 years. *Br Med J* 1998, 317:1181–1184
34. DerSIMONIAN R, CHARETTE LJ, McPEEK B, MOSTELLER F. Reporting on methods in clinical trials. *N Engl J Med* 1982, 306:1332–1337
35. ANONYMOUS. A proposal for structured reporting of randomized controlled trials. The Standards of Reporting Trials Group. *JAMA* 1994, 272:1926–1931
36. ANONYMOUS. Call for comments on a proposal to improve reporting of clinical trials in the biomedical literature. Working Group on Recommendations for Reporting of Clinical Trials in the Biomedical Literature. *Ann Intern Med* 1994, 121:894–895
37. RENNIE D. Reporting randomized controlled trials. An experiment and a call for responses from readers. *JAMA* 1995, 273:1054–1055
38. BEGG C, CHO M, EASTWOOD S, HORTON R, MOHER D, OLKIN I ET AL. Improving the quality of reporting of randomized controlled trials. The CONSORT statement. *JAMA* 1996, 276:637–639
39. PIAGGIO G, ELBOURNE DR, ALTMAN DG, POCKOCK SJ, EVANS SJ, CONSORT GROUP. Reporting of noninferiority and equivalence randomized trials: An extension of the CONSORT statement. *JAMA* 2006, 295:1152–1160
40. CAMPBELL MK, ELBOURNE DR, ALTMAN DG, CONSORT GROUP. CONSORT statement: Extension to cluster randomised trials. *Br Med J* 2004, 328:702–708
41. ZWARENSTEIN M, TREWEEK S, GAGNIER JJ, ALTMAN DG, TUNIS S, HAYNES B ET AL. Improving the reporting of pragmatic trials: An extension of the CONSORT statement. *Br Med J* 2008, 337:a2390
42. IOANNIDIS JP, EVANS SJ, GØTZSCHE PC, O'NEILL RT, ALTMAN DG, SCHULZ K ET AL. Better reporting of harms in randomized trials: An extension of the CONSORT statement. *Ann Intern Med* 2004, 141:781–788
43. BOUTRON I, MOHER D, ALTMAN DG, SCHULZ KF, RAVAUD P, CON-

- SORT GROUP. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: Explanation and elaboration. *Ann Intern Med* 2008, 148:295–309
44. GAGNIER JJ, BOON H, ROCHON P, MOHER D, BARNES J, BOMBARDIER C ET AL. Reporting randomized, controlled trials of herbal interventions: An elaborated CONSORT statement. *Ann Intern Med* 2006, 144:364–367
  45. HOPEWELL S, CLARKE M, MOHER D, WAGER E, MIDDLETON P, ALTMAN DG ET AL. CONSORT for reporting randomized controlled trials in journal and conference abstracts: Explanation and elaboration. *PLoS Med* 2008, 5:e20
  46. SIEGEL JE, WEINSTEIN MC, RUSSELL LB, GOLD MR. Recommendations for reporting cost-effectiveness analyses. Panel on Cost-Effectiveness in Health and Medicine. *JAMA* 1996, 276:1339–1341
  47. DRUMMOND MF, JEFFERSON TO. Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party. *Br Med J* 1996, 313:275–283
  48. LANG TA, SECIC M. *How to report statistics in medicine. Annotated guidelines for authors, editors, and reviewers.* ACP, Philadelphia, 1997
  49. DAVIDOFF F. News from the International Committee of Medical Journal Editors. *Ann Intern Med* 2000, 133:229–231
  50. PLINT AC, MOHER D, MORRISON A, SCHULZ K, ALTMAN DG, HILL C ET AL. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* 2006, 185:263–267
  51. EGGER M, JÜNI P, BARTLETT C, CONSORT GROUP (CONSOLIDATED STANDARDS OF REPORTING OF TRIALS). Value of flow diagrams in reports of randomized controlled trials. *JAMA* 2001, 285:1996–1999
  52. MOHER D, SCHULZ KF, ALTMAN DG, CONSORT GROUP (CONSOLIDATED STANDARDS OF REPORTING OF TRIALS). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *Ann Intern Med* 2001, 134:657–662
  53. MOHER D, SCHULZ KF, ALTMAN D, CONSORT GROUP (CONSOLIDATED STANDARDS OF REPORTING OF TRIALS). The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomized trials. *JAMA* 2001, 285:1987–1991
  54. MOHER D, SCHULZ KF, ALTMAN DG. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet* 2001, 357:1191–1194
  55. CHAN AW, HRÓBJARTSSON A, HAAHR MT, GØTZSCHE PC, ALTMAN DG. Empirical evidence for selective reporting of outcomes in randomized trials: Comparison of protocols to published articles. *JAMA* 2004, 291:2457–2465
  56. AL-MARZOUKI S, ROBERTS I, EVANS S, MARSHALL T. Selective reporting in clinical trials: Analysis of trial protocols accepted by the Lancet. *Lancet* 2008, 372:201
  57. DWAN K, ALTMAN DG, ARNAIZ JA, BLOOM J, CHAN AW, CRONIN E ET AL. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLoS One* 2008, 3:e3081
  58. ALTMAN DG, SCHULZ KF, MOHER D, EGGER M, DAVIDOFF F, ELBOURNE D ET AL. The revised CONSORT statement for reporting randomized trials: Explanation and elaboration. *Ann Intern Med* 2001, 134:663–694
  59. SCHULZ KF, ALTMAN DG, MOHER D, CONSORT GROUP. CONSORT 2010 statement: Updated guidelines for reporting parallel group randomised trials. *Br Med J* 2010, 340:c332
  60. POCOCK SJ. *Clinical trials – a practical approach.* John Wiley, Chichester, NY, 1983
  61. MEINERT CL. *Clinical trials: Design, conduct and analysis.* Oxford University Press, New York, 1986
  62. FRIEDMAN LM, FURBERG CD, DeMETS DL. *Fundamentals of clinical trials.* 3rd ed. Springer-Verlag, New York, 1998
  63. BOLLIGER CT, ZELLWEGER JP, DANIELSSON T, VAN BILJON X, ROBIDOU A, WESTIN A ET AL. Smoking reduction with oral nicotine inhalers: Double blind, randomised clinical trial of efficacy and safety. *Br Med J* 2000, 321:329–333
  64. DICKERSIN K, MANHEIMER E, WIELAND S, ROBINSON KA, LEFEBVRE C, McDONALD S. Development of the Cochrane Collaboration's CENTRAL Register of controlled clinical trials. *Eval Health Prof* 2002, 25:38–64
  65. HOPEWELL S, CLARKE M, MOHER D, WAGER E, MIDDLETON P, ALTMAN DG ET AL. CONSORT for reporting randomised trials in journal and conference abstracts. *Lancet* 2008, 371:281–283
  66. ANONYMOUS. The impact of open access upon public health. *PLoS Med* 2006, 3:e252
  67. HARBOURT AM, KNECHT LS, HUMPHREYS BL. Structured abstracts in MEDLINE, 1989–1991. *Bull Med Libr Assoc* 1995, 83:190–195
  68. HARRIS AH, STANDARD S, BRUNNING JL, CASEY SL, GOLDBERG JH, OLIVER L ET AL. The accuracy of abstracts in psychology journals. *J Psychol* 2002, 136:141–148
  69. PITKIN RM, BRANAGAN MA, BURMEISTER LF. Accuracy of data in abstracts of published research articles. *JAMA* 1999, 281:1110–1111
  70. WARD LG, KENDRACH MG, PRICE SO. Accuracy of abstracts for original research articles in pharmacy journals. *Ann Pharmacother* 2004, 38:1173–1177
  71. GØTZSCHE PC. Believability of relative risks and odds ratios in abstracts: Cross sectional study. *Br Med J* 2006, 333:231–234
  72. IOANNIDIS JP, LAU J. Completeness of safety reporting in randomized trials: An evaluation of 7 medical areas. *JAMA* 2001, 285:437–443
  73. HAYNES RB, MULROW CD, HUTH EJ, ALTMAN DG, GARDNER MJ. More informative abstracts revisited. *Ann Intern Med* 1990, 113:69–76
  74. TADDIO A, PAIN T, FASSOS FF, BOON H, ILLERSICH AL, EINARSON TR. Quality of nonstructured and structured abstracts of original research articles in the British Medical Journal, the Canadian Medical Association Journal and the Journal of the American Medical Association. *CMAJ* 1994, 150:1611–1615
  75. WAGER E, MIDDLETON P. Technical editing of research reports in biomedical journals. *Cochrane Database Syst Rev* 2008, 4:MR000002
  76. HARTLEY J, SYDES M, BLURTON A. Obtaining information accurately and quickly: Are structured abstracts more efficient? *J Inform Sci* 1996, 22:349–356
  77. GILLIGAN D, NICOLSON M, SMITH I, GROENH, DALESIO O, GOLDSTRAW P ET AL. Preoperative chemotherapy in patients with respectable non-small cell lung cancer: Results of the MRC LU22/NVALT2/EORTC 08012 multicentre randomised trial and up-

- date of systematic review. *Lancet* 2007, 369:1929–1937
78. SANDLER AD, SUTTON KA, DeWEESE J, GIRARDI MA, SHEPPARD V, BODFISH JW. Lack of benefit of a single dose of synthetic human secretin in the treatment of autism and pervasive developmental disorder. *N Engl J Med* 1999, 341:1801–1806
  79. WORLD MEDICAL ASSOCIATION. Declaration of Helsinki: Ethical principles for medical research involving human subjects. 59th WMA General Assembly, Seoul 2008. Available at: [www.wma.net/e/policy/b3.htm](http://www.wma.net/e/policy/b3.htm) (accessed 2.6.2009)
  80. LAU J, ANTMAN EM, JIMENEZ-SILVA J, KUPELNICK B, MOSTELLER F, CHALMERS TC. Cumulative meta-analysis of therapeutic trials for myocardial infarction. *N Engl J Med* 1992, 327:248–254
  81. FERGUSSON D, GLASS KC, HUTTON B, SHAPIRO S. Randomized controlled trials of aprotinin in cardiac surgery: Could clinical equipoise have stopped the bleeding? *Clin Trials* 2005, 2:218–229
  82. SAVULESCU J, CHALMERS I, BLUNT J. Are research ethics committees behaving unethically? Some suggestions for improving performance and accountability. *Br Med J* 1996, 313:1390–1393
  83. SADLER LC, DAVISON T, McCOWAN LM. A randomised controlled trial and meta-analysis of active management of labour. *BJOG* 2000, 107:909–915
  84. BATH FJ, OWEN VE, BATH PM. Quality of full and final publications reporting acute stroke trials: A systematic review. *Stroke* 1998, 29:2203–2210
  85. BLUMER JL, FINDLING RL, SHIH WJ, SOUBRANE C, REED MD. Controlled clinical trial of zolpidem for the treatment of insomnia associated with attention-deficit/hyperactivity disorder in children 6 to 17 years of age. *Pediatrics* 2009, 123:e770–e776
  86. SABATINE MS, ANTMAN EM, WIDIMSKY P, EBRAHIM IO, KISS RG, SAAIMAN A ET AL. Otamixaban for the treatment of patients with non-ST-elevation acute coronary syndromes (SEPIA-ACS1 TIMI 42): A randomised, double-blind, active-controlled, phase 2 trial. *Lancet* 2009, 374:787–795
  87. GRANT AM, ALTMAN DG, BABIKER AB, CAMPBELL MK, CLEMENS FJ, DARBYSHIRE JH ET AL. Issues in data monitoring and interim analysis of trials. *Health Technol Assess* 2005, 9:1–238, iii–iv
  88. GALLO P, KRAMS M. PhRMA Working Group on adaptive designs: Full white paper. *Drug Information Journal* 2006, 40:421–483
  89. BROWN CH, TEN HAVE TR, JO B, DAGNE G, WYMAN PA, MUTHÉN B ET AL. Adaptive designs for randomized trials in public health. *Annu Rev Public Health* 2009, 30:1–25
  90. KELLY PJ, SOORIYARACHCHI MR, STALLARD N, TODD S. A practical comparison of group-sequential and adaptive designs. *J Biopharm Stat* 2005, 15:719–738
  91. PILDAL J, CHAN AW, HRÖBJARTSSON A, FORFANG E, ALTMAN DG, GØTZSCHE PC. Comparison of descriptions of allocation concealment in trial protocols and the published reports: Cohort study. *Br Med J* 2005, 330:1049
  92. CHAN AW, HRÖBJARTSSON A, JØRGENSEN KJ, GØTZSCHE PC, ALTMAN DG. Discrepancies in sample size calculations and data analyses reported in randomised trials: Comparison of publications with protocols. *Br Med J* 2008, 337:a2299
  93. NDEKHA MJ, VAN OOSTERHOUT JJ, ZIJLSTRA EE, MANARY M, SALLOOJEE H, MANARY MJ. Supplementary feeding with either ready-to-use fortified spread or corn-soy blend in wasted adults starting antiretroviral therapy in Malawi: Randomised, investigator blinded, controlled trial. *Br Med J* 2009, 338:b1867
  94. ROTHWELL PM. External validity of randomised controlled trials: “To whom do the results of this trial apply?”. *Lancet* 2005, 365:82–93
  95. FUKS A, WEIJER C, FREEDMAN B, SHAPIRO S, SKRUTKOWSKA M, RIAZ A. A study in contrasts: Eligibility criteria in a twenty-year sample of NSABP and POG clinical trials. National Surgical Adjuvant Breast and Bowel Program. Pediatric Oncology Group. *J Clin Epidemiol* 1998, 51:69–79
  96. SHAPIRO SH, WEIJER C, FREEDMAN B. Reporting the study populations of clinical trials. Clear transmission or static on the line? *J Clin Epidemiol* 2000, 53:973–979
  97. GANDHI M, AMELI N, BACCHETTI P, SHARP GB, FRENCH AL, YOUNG M ET AL. Eligibility criteria for HIV clinical trials and generalizability of results: The gap between published reports and study protocols. *AIDS* 2005, 19:1885–1896
  98. HALL JC, MILLS B, NGUYEN H, HALL JL. Methodologic standards in surgical trials. *Surgery* 1996, 119:466–472
  99. WEISS NS, KOEPESELL TD, PSATY BM. Generalizability of the results of randomized trials. *Arch Intern Med* 2008, 168:133–135
  100. POISE STUDY GROUP, DEVEREAUX PJ, YANG H, YUSUF S, GUYATT G, LESLIE K ET AL. Effects of extended-release metoprolol succinate in patients undergoing non-cardiac surgery (POISE trial): A randomised controlled trial. *Lancet* 2008, 371:1839–1847
  101. RANNOU F, DIMET J, BOUTRON I, BARON G, FAYAD F, MACÉ Y ET AL. Splint for base-of-thumb osteoarthritis: A randomized trial. *Ann Intern Med* 2009, 150:661–669
  102. GLASZIOU P, MEATS E, HENEGHAN C, SHEPPERD S. What is missing from descriptions of treatment in trials and reviews? *Br Med J* 2008, 336:1472–1474
  103. MEASE PJ, GOFFE BS, METZ J, VANDERSTOEP A, FINCK B, BURGE DJ. Etanercept in the treatment of psoriatic arthritis and psoriasis: A randomised trial. *Lancet* 2000, 356:385–390
  104. McDOWELL I, NEWELL C. *Measuring health: A guide to rating scales and questionnaires*. 3rd ed. Oxford University Press, New York, 2006
  105. STREINER DL, NORMAN GR. *Health measurement scales: A practical guide to their development and use*. 3rd ed. Oxford University Press, Oxford, 2003
  106. CLARKE M. Standardising outcomes for clinical trials and systematic reviews. *Trials* 2007, 8:39
  107. SANDERS C, EGGER M, DONOVAN J, TALLON D, FRANKEL S. Reporting on quality of life in randomised controlled trials: Bibliographic study. *Br Med J* 1998, 317:1191–1194
  108. GØTZSCHE PC. Methodology and overt and hidden bias in reports of 196 double-blind trials of nonsteroidal antiinflammatory drugs in rheumatoid arthritis. *Control Clin Trials* 1989, 10:31–56
  109. MARSHALL M, LOCKWOOD A, BRADLEY C, ADAMS C, JOY C, FENTON M. Unpublished rating scales: A major source of bias in randomised controlled trials of treatments for schizophrenia. *Br J Psychiatry* 2000, 176:249–252
  110. JADAD AR, BOYLE M, CUNNINGHAM C, KIM M, SCHACHAR R. Treatment of attention-deficit/hyperactivity disorder. Evidence report/Technology assessment no 11. US Department of Health and Human Services, Public Health Service, Agency for Healthcare Research and Quality. AHQR publication no 00-E005, Rockville, MD, 1999



111. SCHACHTER HM, PHAM B, KING J, LANGFORD S, MOHER D. The efficacy and safety of methylphenidate in attention deficit disorder: A systematic review and meta-analysis. Prepared for the Therapeutics Initiative. Vancouver, BC, and the British Columbia Ministry for Children and Families, 2000
112. DARGIE HJ. Effect of carvedilol on outcome after myocardial infarction in patients with left-ventricular dysfunction: The CAPRICORN randomised trial. *Lancet* 2001, 357:1385–1390
113. CHAN AW, KRLEZA-JERIC K, SCHMID I, ALTMAN DG. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ* 2004, 171:735–740
114. VERMEULEN H, HOFLAND J, LEGEMATE DA, UBBINK DT. Intravenous fluid restriction after major abdominal surgery: A randomized blinded clinical trial. *Trials* 2009, 10:50
115. FOX K, FORD I, STEG PG, TENDERA M, FERRARI R, BEAUTIFUL INVESTIGATORS. Ivabradine for patients with stable coronary artery disease and left-ventricular systolic dysfunction (BEAUTIFUL): A randomised, double-blind, placebo-controlled trial. *Lancet* 2008, 372:807–816
116. CAMPBELL MJ, JULIOUS SA, ALTMAN DG. Estimating sample sizes for binary, ordered categorical, and continuous outcomes in two group comparisons. *Br Med J* 1995, 311:1145–1148
117. GUYATT GH, MILLS EJ, ELBOURNE D. In the era of systematic reviews, does the size of an individual trial still matter. *PLoS Med* 2008, 5:e4
118. SCHULZ KF, GRIMES DA. Sample size calculations in randomised trials: Mandatory and mystical. *Lancet* 2005, 365:1348–1353
119. HALPERN SD, KARLAWISH JH, BERLIN JA. The continuing unethical conduct of underpowered clinical trials. *JAMA* 2002, 288:358–362
120. ALTMAN DG, BLAND JM. Absence of evidence is not evidence of absence. *Br Med J* 1995, 311:485
121. MOHER D, DULBERG CS, WELLS GA. Statistical power, sample size, and their reporting in randomized controlled trials. *JAMA* 1994, 272:122–124
122. FREIMAN JA, CHALMERS TC, SMITH H Jr, KUEBLER RR. The importance of beta, the type II error and sample size in the design and interpretation of the randomized control trial. Survey of 71 “negative” trials. *N Engl J Med* 1978, 299:690–694
123. CHARLES P, GIRAudeau B, DECHARTRES A, BARON G, RAVAUD P. Reporting of sample size calculation in randomised controlled trials: Review. *Br Med J* 2009, 338:b1732
124. YUSUF S, COLLINS R, PETO R. Why do we need some large, simple randomized trials? *Stat Med* 1984, 3:409–422
125. GOODMAN SN, BERLIN JA. The use of predicted confidence intervals when planning experiments and the misuse of power when interpreting results. *Ann Intern Med* 1994, 121:200–206
126. GALGIANI JN, CATANZARO A, CLOUD GA, JOHNSON RH, WILLIAMS PL, MIRELS LF ET AL. Comparison of oral fluconazole and itraconazole for progressive, nonmeningeal coccidioidomycosis. A randomized, double-blind trial. Mycoses Study Group. *Ann Intern Med* 2000, 133:676–686
127. ACTIVE INVESTIGATORS, CONNOLLY SJ, POGUE J, HART RG, HOHNLOSER SH, PFEFFER M ET AL. Effect of clopidogrel added to aspirin in patients with atrial fibrillation. *N Engl J Med* 2009, 360:2066–2078
128. GELLER NL, POCOCK SJ. Interim analyses in randomized clinical trials: Ramifications and guidelines for practitioners. *Biometrics* 1987, 43:213–223
129. BERRY DA. Interim analyses in clinical trials: Classical vs Bayesian approaches. *Stat Med* 1985, 4:521–526
130. POCOCK SJ. When to stop a clinical trial. *Br Med J* 1992, 305:235–240
131. DeMETS DL, POCOCK SJ, JULIAN DG. The agonising negative trend in monitoring of clinical trials. *Lancet* 1999, 354:1983–1988
132. BUYSE M. Interim analyses, stopping rules and data monitoring in clinical trials in Europe. *Stat Med* 1993, 12:509–520
133. SYDES MR, ALTMAN DG, BABIKER AB, PARMAR MK, SPIEGELHALTER DJ, DAMOCLES GROUP. Reported use of data monitoring committees in the main published reports of randomized controlled trials: A cross-sectional study. *Clin Trials* 2004, 1:48–59
134. MONTORI VM, DEVEREAUX PJ, ADHIKARI NK, BURNS KE, EGGERT CH, BRIEL M ET AL. Randomized trials stopped early for benefit: A systematic review. *JAMA* 2005, 294:2203–2209
135. COUTINHO IC, RAMOS DE, AMORIM MM, KATZ L, BANDEIRA DE, FERRAZ AA. Uterine exteriorization compared with *in situ* repair at caesarean delivery: A randomized controlled trial. *Obstet Gynecol* 2008, 111:639–647
136. JÜNI P, ALTMAN DG, EGGER M. Assessing the quality of controlled clinical trials. In: Egger M, Davey Smith G, Altman DG (eds) *Systematic reviews in health care: Meta-analysis in context*. BMJ Books, London, 2001
137. CREININ MD, MEYN LA, BORGATTA L, BARNHART K, JENSEN J, BURKE AE ET AL. Multicenter comparison of the contraceptive ring and patch: A randomized controlled trial. *Obstet Gynecol* 2008, 111:267–277
138. TATE DF, JACKVONY EH, WING RR. Effects of internet behavioral counseling on weight loss in adults at risk for type 2 diabetes: A randomized trial. *JAMA* 2003, 289:1833–1836
139. LACHIN JM. Properties of simple randomization in clinical trials. *Control Clin Trials* 1988, 9:312–326
140. PETO R, PIKE MC, ARMITAGE P, BRESLOW NE, COX DR, HOWARD SV ET AL. Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *Br J Cancer* 1976, 34:585–612
141. SCHULZ KF, GRIMES DA. *The Lancet: Handbook of essential concepts in clinical research*. Elsevier, London, 2006
142. ALTMAN DG, BLAND JM. How to randomise. *Br Med J* 1999, 319:703–704
143. SCHULZ KF. Subverting randomization in controlled trials. *JAMA* 1995, 274:1456–1458
144. ENAS GG, ENAS NH, SPRADLIN CT, WILSON MG, WILTSE CG. Baseline comparability in clinical trials: Prevention of poststudy anxiety. *Drug Information Journal* 1990, 24:541–548
145. TREASURE T, MACRAE KD. Minimisation: the platinum standard for trials? Randomisation doesn’t guarantee similarity of groups; minimisation does. *Br Med J* 1998, 317:362–363
146. SINEI SK, SCHULZ KF, LAMPTEY PR, GRIMES DA, MATI JK, ROSENTHAL SM ET AL. Preventing IUCD-related pelvic infection: The efficacy of prophylactic doxycycline at insertion. *Br J Obstet Gynaecol* 1990, 97:412–419
147. RADFORD JA, LANDORF KB, BUCHBINDER R, COOK C. Effectiveness of low-Dye taping for the short-term treatment of plantar heel pain: A randomised trial. *BMC Musculoskelet Disord*



- 2006, 7:64
148. CHALMERS TC, LEVIN H, SACKS HS, REITMAN D, BERRIER J, NAGALINGAM R. Meta-analysis of clinical trials as a scientific discipline. I: Control of bias and comparison with large co-operative trials. *Stat Med* 1987, 6:315–328
  149. POCOCK SJ. Statistical aspects of clinical trial design. *Statisticalian* 1982, 31:1–18
  150. HAAG U. Technologies for automating randomized treatment assignment in clinical trials. *Drug Information Journal* 1998, 32:11
  151. PIAGGIO G, ELBOURNE D, SCHULZ KF, VILLAR J, PINOL AP, GÜLMEZOGLU AM ET AL. The reporting of methods for reducing and detecting bias: An example from the WHO Misoprostol Third Stage of Labour equivalence randomised controlled trial. *BMC Med Res Methodol* 2003, 3:19
  152. PILDAL J, HRÖBJARTSSON A, JØRGENSEN KJ, HILDEN J, ALTMAN DG, GØTZSCHE PC. Impact of allocation concealment on conclusions drawn from meta-analyses of randomized trials. *Int J Epidemiol* 2007, 36:847–857
  153. WOOD L, EGGER M, GLUUD LL, SCHULZ KF, JÜNI P, ALTMAN DG ET AL. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: Meta-epidemiological study. *Br Med J* 2008, 336:601–605
  154. MCCANDLISH R, BOWLER U, VAN ASTEN H, BERRIDGE G, WINTER C, SAMES L ET AL. A randomised controlled trial of care of the perineum during second stage of normal labour. *Br J Obstet Gynaecol* 1998, 105:1262–1272
  155. WEBSTER J, CLARKE S, PATERSON D, HUTTON A, VAN DYK S, GALE C ET AL. Routine care of peripheral intravenous catheters versus clinically indicated replacement: Randomised controlled trial. *Br Med J* 2008, 337:a339
  156. SMITH SA, SHAH ND, BRYANT SC, CHRISTIANSON TJ, BJORNSEN SS, GIESLER PD ET AL. Chronic care model and shared care in diabetes: Randomized trial of an electronic decision support system. *Mayo Clin Proc* 2008, 83:747–757
  157. SACKS FM, BRAY GA, CAREY VJ, SMITH SR, RYAN DH, ANTON SD ET AL. Comparison of weight-loss diets with different compositions of fat, protein, and carbohydrates. *N Engl J Med* 2009, 360:859–873
  158. KAPTCHUK TJ. Intentional ignorance: A history of blind assessment and placebo controls in medicine. *Bull Hist Med* 1998, 72:389–433
  159. GUYATT GH, PUGSLEY SO, SULLIVAN MJ, THOMPSON PJ, BERMAN L, JONES NL ET AL. Effect of encouragement on walking test performance. *Thorax* 1984, 39:818–822
  160. GØTZSCHE PC. Blinding during data analysis and writing of manuscripts. *Control Clin Trials* 1996, 17:285–290
  161. KARLOWSKI TR, CHALMERS TC, FRENKEL LD, KAPIKIAN AZ, LEWIS TL, LYNCH JM. Ascorbic acid for the common cold. A prophylactic and therapeutic trial. *JAMA* 1975, 231:1038–1042
  162. NOSEWORTHY JH, EBERS GC, VANDERVOORT MK, FARQUHAR RE, YETISIR E, ROBERTS R. The impact of blinding on the results of a randomized, placebo-controlled multiple sclerosis clinical trial. *Neurology* 1994, 44:16–20
  163. CARLEY SD, LIBETTA C, FLAVIN B, BUTLER J, TONG N, SAMMY I. An open prospective randomised trial to reduce the pain of blood glucose testing: Ear versus thumb. *Br Med J* 2000, 321:20
  164. SCHULZ KF, CHALMERS I, ALTMAN DG. The landscape and lexicon of blinding in randomized trials. *Ann Intern Med* 2002, 136:254–259
  165. DAY SJ, ALTMAN DG. Statistics notes: Blinding in clinical trials and other studies. *Br Med J* 2000, 321:504
  166. MONTORI VM, BHANDARI M, DEVEREAUX PJ, MANNS BJ, GHALI WA, GUYATT GH. In the dark: The reporting of blinding status in randomized controlled trials. *J Clin Epidemiol* 2002, 55:787–790
  167. CHENG K, SMYTH RL, MOTLEY J, O’HEA U, ASHBY D. Randomized controlled trials in cystic fibrosis (1966–1997) categorized by time, design, and intervention. *Pediatr Pulmonol* 2000, 29:1–7
  168. LANG T. Masking or blinding? An unscientific survey of mostly medical journal editors on the great debate. *MedGenMed* 2000, 2:E25
  169. DEVEREAUX PJ, MANNS BJ, GHALI WA, QUAN H, LACCHETTI C, MONTORI VM ET AL. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 2001, 285:2000–2003
  170. HAAHR MT, HRÖBJARTSSON A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. *Clin Trials* 2006, 3:360–365
  171. MEINERT CL. Masked monitoring in clinical trials – blind stupidity? *N Engl J Med* 1998, 338:1381–1382
  172. MILLS E, PROUSKY J, RASKIN G, GAGNIER J, RACHLIS B, MONTORI VM ET AL. The safety of over-the-counter niacin. A randomized placebo-controlled trial [ISRCTN18054903]. *BMC Clin Pharmacol* 2003, 3:4
  173. SCHULZ KF, GRIMES DA, ALTMAN DG, HAYES RJ. Blinding and exclusions after allocation in randomised controlled trials: Survey of published parallel group trials in obstetrics and gynaecology. *Br Med J* 1996, 312:742–744
  174. FERGUSSON D, GLASS KC, WARING D, SHAPIRO S. Turning a blind eye: The success of blinding reported in a random sample of randomised, placebo controlled trials. *Br Med J* 2004, 328:432
  175. SACKETT DL. Turning a blind eye: Why we don’t test for blindness at the end of our trials. *Br Med J* 2004, 328:1136
  176. ASTRUP A, RÖSSNER S, VAN GAAL L, RISSANEN A, NISKANEN L, AL HAKIM M ET AL. Effects of liraglutide in the treatment of obesity: A randomised, double-blind, placebo-controlled study. *Lancet* 2009, 374:1606–1616
  177. ALTMAN DG, GORE SM, GARDNER MJ, POCOCK SJ. Statistical guidelines for contributors to medical journals. In: Altman DG, Machin D, Bryant TN, Gardner MJ (eds) *Statistics with confidence: Confidence intervals and statistical guidelines*. 2nd ed. BMJ Books, London, 2000:171–190
  178. ALTMAN DG, BLAND JM. Statistics notes. Units of analysis. *Br Med J* 1997, 314:1874
  179. BOLTON S. Independence and statistical inference in clinical trial designs: A tutorial review. *J Clin Pharmacol* 1998, 38:408–412
  180. GREENLAND S. Principles of multilevel modelling. *Int J Epidemiol* 2000, 29:158–167
  181. ALBERT CM, COOK NR, GAZIANO JM, ZAHARRIS E, MACFADYEN J, DANIELSON E ET AL. Effect of folic acid and B vitamins on risk

- of cardiovascular events and total mortality among women at high risk for cardiovascular disease: A randomized trial. *JAMA* 2008, 299:2027–2036
182. MATTHEWS JN, ALTMAN DG. Interaction 3: How to examine heterogeneity. *Br Med J* 1996, 313:862
  183. ASSMANN SF, POCOCK SJ, ENOS LE, KASTEN LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 2000, 355:1064–1069
  184. MATTHEWS JN, ALTMAN DG. Statistics notes. Interaction 2: Compare effect sizes not P values. *Br Med J* 1996, 313:808
  185. OXMAN AD, GUYATT GH. A consumer's guide to subgroup analyses. *Ann Intern Med* 1992, 116:78–84
  186. STEYERBERG EW, BOSSUYT PM, LEE KL. Clinical trials in acute myocardial infarction: Should we adjust for baseline characteristics? *Am Heart J* 2000, 139:745–751
  187. ALTMAN DG. Adjustment for covariate imbalance. In: Armitage P, Colton T (eds) *Encyclopedia of biostatistics*. John Wiley, New York, 1998:1000–1005
  188. MÜLLNER M, MATTHEWS H, ALTMAN DG. Reporting on statistical methods to adjust for confounding: A cross-sectional survey. *Ann Intern Med* 2002, 136:122–126
  189. CONCATO J, FEINSTEIN AR, HOLFORD TR. The risk of determining risk with multivariable models. *Ann Intern Med* 1993, 118:201–210
  190. BENDER R, GROUVEN U. Logistic regression models used in medical research are poorly presented. *Br Med J* 1996, 313:628
  191. KHAN KS, CHIEN PF, DWARAKANATH LS. Logistic regression models in obstetrics and gynecology literature. *Obstet Gynecol* 1999, 93:1014–1020
  192. SACKETT DL, GENT M. Controversy in counting and attributing events in clinical trials. *N Engl J Med* 1979, 301:1410–1412
  193. MAY GS, DeMETS DL, FRIEDMAN LM, FURBERG C, PASSAMANI E. The randomized clinical trial: Bias in analysis. *Circulation* 1981, 64:669–673
  194. ALTMAN DG, CUZICK J, PETO J. More on zidovudine in asymptomatic HIV infection. *N Engl J Med* 1994, 330:1758–1759
  195. MEINERT CL. Beyond CONSORT: Need for improved reporting standards for clinical trials. Consolidated Standards of Reporting Trials. *JAMA* 1998, 279:1487–1489
  196. GRANT AM, WILEMAN SM, RAMSAY CR, MOWAT NA, KRUKOWSKI ZH, HEADING RC ET AL. Minimal access surgery compared with medical management for chronic gastro-oesophageal reflux disease: UK collaborative randomised trial. *Br Med J* 2008, 337:a2664
  197. VAN LOON AJ, MANTINGH A, SERLIER EK, KROON G, MOOYAART EL, HUISJES HJ. Randomised controlled trial of magnetic-resonance pelvimetry in breech presentation at term. *Lancet* 1997, 350:1799–1804
  198. BROWN MJ, PALMER CR, CASTAIGNE A, DE LEEUW PW, MANCIA G, ROSENTHAL T ET AL. Morbidity and mortality in patients randomised to double-blind treatment with a long-acting calcium-channel blocker or diuretic in the International Nifedipine GITS study: Intervention as a Goal in Hypertension Treatment (INSIGHT). *Lancet* 2000, 356:366–372
  199. LaCROIX AZ, OTT SM, ICHIKAWA L, SCHOLES D, BARLOW WE. Low-dose hydrochlorothiazide and preservation of bone mineral density in older adults. A randomized, double-blind, placebo-controlled trial. *Ann Intern Med* 2000, 133:516–526
  200. SHUSTER JJ. Median follow-up in clinical trials. *J Clin Oncol* 1991, 9:191–192
  201. ALTMAN DG, DE STAVOLA BL, LOVE SB, STEPNIIEWSKA KA. Review of survival analyses published in cancer journals. *Br J Cancer* 1995, 72:511–518
  202. AUVERT B, TALJAARD D, LAGARDE E, SOBNGWI-TAMBEKOU J, SITA R, PUREN A. Randomized, controlled intervention trial of male circumcision for reduction of HIV infection risk: The ANRS 1265 Trial. *PLoS Med* 2005, 2:e298
  203. DIGGLE L, DEEKS J. Effect of needle length on incidence of local reactions to routine immunisation in infants aged 4 months: Randomised controlled trial. *Br Med J* 2000, 321:931–933
  204. POCOCK S, WHITE I. Trials stopped early: Too good to be true? *Lancet* 1999, 353:943–944
  205. HUGHES MD, POCOCK SJ. Stopping rules and estimation problems in clinical trials. *Stat Med* 1988, 7:1231–1242
  206. KIRI A, TONASCIA S, MEINERT CL. Treatment effects monitoring committees and early stopping in large clinical trials. *Clin Trials* 2004, 1:40–47
  207. PSATY BM, RENNIE D. Stopping medical research to save money: A broken pact with researchers and patients. *JAMA* 2003, 289:2128–2131
  208. TEMPLE R. FDA perspective on trials with interim efficacy evaluations. *Stat Med* 2006, 25:3245–3249
  209. TELMISARTAN RANDOMISED ASSESSMENT STUDY IN ACE INTOLERANT SUBJECTS WITH CARDIOVASCULAR DISEASE (TRANSCEND) INVESTIGATORS, YUSUF S, TEO K, ANDERSON C, POGUE J, DYAL L ET AL. Effects of the angiotensin-receptor blocker telmisartan on cardiovascular events in high-risk patients intolerant to angiotensin-converting enzyme inhibitors: A randomised controlled trial. *Lancet* 2008, 372:1174–1183
  210. SENN S. Base logic: Tests of baseline balance in randomized clinical trials. *Clin Res Regulatory Affairs* 1995, 12:171–182
  211. ALTMAN DG. Comparability of randomised groups. *Statistician* 1985, 34:125–136
  212. HEIT JA, ELLIOTT CG, TROWBRIDGE AA, MORREY BF, GENT M, HIRSH J. Ardeparin sodium for extended out-of-hospital prophylaxis against venous thromboembolism after total hip or knee replacement. A randomized, double-blind, placebo-controlled trial. *Ann Intern Med* 2000, 132:853–861
  213. HADERSLEV KV, TJELLESEN L, SORENSEN HA, STAUN M. Alendronate increases lumbar spine bone mineral density in patients with Crohn's disease. *Gastroenterology* 2000, 119:639–646
  214. FIELDS WS, MASLENIKOV V, MEYER JS, HASS WK, REMINGTON RD, MACDONALD M. Joint study of extracranial arterial occlusion. V. Progress report of prognosis following surgery or non-surgical treatment for transient cerebral ischemic attacks and cervical carotid artery lesions. *JAMA* 1970, 211:1993–2003
  215. LEE YJ, ELLENBERG JH, HIRTZ DG, NELSON KB. Analysis of clinical trials by treatment actually received: Is it really an option? *Stat Med* 1991, 10:1595–1605
  216. LEWIS JA, MACHIN D. Intention to treat – who should use ITT? *Br J Cancer* 1993, 68:647–650
  217. LACHIN JL. Statistical considerations in the intent-to-treat principle. *Control Clin Trials* 2000, 21:526
  218. SHEINER LB, RUBIN DB. Intention-to-treat analysis and the goals

- of clinical trials. *Clin Pharmacol Ther* 1995, 57:6–15
219. NAGELKERKE N, FIDLER V, BERNSEN R, BORGDORFF M. Estimating treatment effects in randomized clinical trials in the presence of non-compliance. *Stat Med* 2000, 19:1849–1864
220. MELANDER H, AHLQVIST-RASTAD J, MEIJER G, BEERMANN B. Evidence b(i)ased medicine – selective reporting from studies sponsored by pharmaceutical industry: Review of studies in new drug applications. *Br Med J* 2003, 326:1171–1173
221. GRAVEL J, OPATRYN L, SHAPIRO S. The intention-to-treat approach in randomized controlled trials: Are authors saying what they do and doing what they say? *Clin Trials* 2007, 4:350–356
222. KRUSE RL, ALPER BS, REUST C, STEVERMER JJ, SHANNON S, WILIAMS RH. Intention-to-treat analysis: Who is in? Who is out? *J Fam Pract* 2002, 51:969–971
223. HERMAN A, BOTSER IB, TENENBAUM S, CHECHICK A. Intention-to-treat analysis and accounting for missing data in orthopaedic randomized clinical trials. *J Bone Joint Surg Am* 2009, 91:2137–2143
224. RUIZ-CANELA M, MARTÍNEZ-GONZÁLEZ MA, DE IRALA-ESTÉVEZ J. Intention to treat analysis is related to methodological quality. *Br Med J* 2000, 320:1007–1008
225. ALTMAN DG. Missing outcomes in randomised trials: Addressing the dilemma. *Open Med* 2009, 3:e21–e23
226. WOOD AM, WHITE IR, THOMPSON SG. Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clin Trials* 2004, 1:368–376
227. STREINER DL. Missing data and the trouble with LOCF. *Evid Based Ment Health* 2008, 11:3–5
228. MOLNAR FJ, HUTTON B, FERGUSSON D. Does analysis using “last observation carried forward” introduce bias in dementia research? *CMAJ* 2008, 179:751–753
229. WARE JH. Interpreting incomplete data in studies of diet and weight loss. *N Engl J Med* 2003, 348:2136–2137
230. STREINER DL. The case of the missing data: Methods of dealing with dropouts and other research vagaries. *Can J Psychiatry* 2002, 47:68–75
231. LANE P. Handling drop-out in longitudinal clinical trials: A comparison of the LOCF and MMRM approaches. *Pharm Stat* 2008, 7:93–106
232. ABRAHA I, MONTEODORI A, ROMAGNOLI C. Modified intention to treat: Frequency, definition and implication for clinical trials. XV Cochrane Colloquium, Sao Paulo, Brazil, 2007:86–87 [abstract]
233. ALTMAN DG. Confidence intervals in practice. In: Altman DG, Machin D, Bryant TN, Gardner MJ (eds) *Statistics with confidence*. 2nd ed. BMJ Books, London, 2000:6–14
234. VAN LINSCHOTEN R, VAN MIDDELKOOP M, BERGER MY, HEINTJES EM, VERHAAR JA, WILLEMSSEN SP ET AL. Supervised exercise therapy versus usual care for patellofemoral pain syndrome: An open label randomised controlled trial. *Br Med J* 2009, 339:b4074
235. ALTMAN DG. Clinical trials and meta-analyses. In: Altman DG, Machin D, Bryant TN, Gardner MJ (eds) *Statistics with confidence*. 2nd ed. BMJ Books, London, 2000:120–138
236. ANONYMOUS. Uniform requirements for manuscripts submitted to biomedical journals. International Committee of Medical Journal Editors. *Ann Intern Med* 1997, 126:36–47
237. GARDNER MJ, ALTMAN DG. Confidence intervals rather than P values: Estimation rather than hypothesis testing. *Br Med J (Clin Res Ed)* 1986, 292:746–750
238. BAILAR JC 3rd, MOSTELLER F. Guidelines for statistical reporting in articles for medical journals. Amplifications and explanations. *Ann Intern Med* 1988, 108:266–273
239. BLAND JM. Quoting intermediate analyses can only mislead. *Br Med J* 1997, 314:1907–1908
240. COOK RJ, SACKETT DL. The number needed to treat: A clinically useful measure of treatment effect. *Br Med J* 1995, 310:452–454
241. ALTMAN DG, ANDERSEN PK. Calculating the number needed to treat for trials where the outcome is time to an event. *Br Med J* 1999, 319:1492–1495
242. ANONYMOUS. Early versus delayed neonatal administration of a synthetic surfactant – the judgment of OSIRIS. The OSIRIS Collaborative Group (open study of infants at high risk of or with respiratory insufficiency – the role of surfactant). *Lancet* 1992, 340:1363–1369
243. SORENSEN L, GYRD-HANSEN D, KRISTIANSEN IS, NEXØE J, NIELSEN JB. Laypersons’ understanding of relative risk reductions: Randomised cross-sectional study. *BMC Med Inform Decis Mak* 2008, 8:31
244. BOBBIO M, DEMICHELIS B, GIUSTETTO G. Completeness of reporting trial results: Effect on physicians’ willingness to prescribe. *Lancet* 1994, 343:1209–1211
245. NAYLOR CD, CHEN E, STRAUSS B. Measured enthusiasm: Does the method of reporting trial results alter perceptions of therapeutic effectiveness? *Ann Intern Med* 1992, 117:916–921
246. TUKEY JW. Some thoughts on clinical trials, especially problems of multiplicity. *Science* 1977, 198:679–684
247. YUSUF S, WITTES J, PROBSTFIELD J, TYROLER HA. Analysis and interpretation of treatment effects in subgroups of patients in randomized clinical trials. *JAMA* 1991, 266:93–98
248. HAHN S, WILLIAMSON PR, HUTTON JL, GARNER P, FLYNN EV. Assessing the potential for bias in meta-analysis due to selective reporting of subgroup analyses within studies. *Stat Med* 2000, 19:3325–3336
249. BHANDARI M, DEVEREAUX PJ, LI P, MAH D, LIM K, SCHÜNEMANN HJ ET AL. Misuse of baseline comparison tests and subgroup analyses in surgical trials. *Clin Orthop Relat Res* 2006, 447:247–251
250. LEVIN M, QUINT PA, GOLDSTEIN B, BARTON P, BRADLEY JS, SHERMIE SD ET AL. Recombinant bactericidal/permeability-increasing protein (rBPI21) as adjunctive treatment for children with severe meningococcal sepsis: A randomised trial. rBPI21 Meningococcal Sepsis Study Group. *Lancet* 2000, 356:961–967
251. SCHARF O, COLEVAS AD. Adverse event reporting in publications compared with sponsor database for cancer clinical trials. *J Clin Oncol* 2006, 24:3933–3938
252. PITROU I, BOUTRON I, AHMAD N, RAVAUD P. Reporting of safety results in published reports of randomized controlled trials. *Arch Intern Med* 2009, 169:1756–1761
253. BODEN WE, O’ROURKE RA, TEO KK, HARTIGAN PM, MARON DJ, KOSTUK WJ ET AL. Optimal medical therapy with or without PCI for

- stable coronary disease. *N Engl J Med* 2007, 356:1503–1516
254. HORTON R. The rhetoric of research. *Br Med J* 1995, 310:985–987
  255. ANNALS OF INTERNAL MEDICINE. Information for authors. Available at [www.annals.org](http://www.annals.org) (accessed 15.1.2008)
  256. DOCHERTY M, SMITH R. The case for structuring the discussion of scientific papers. *Br Med J* 1999, 318:1224–1225
  257. PURCELL GP, DONOVAN SL, DAVIDOFF F. Changes to manuscripts during the editorial process: Characterizing the evolution of a clinical paper. *JAMA* 1998, 280:227–228
  258. IOANNIDIS JP. Limitations are not properly acknowledged in the scientific literature. *J Clin Epidemiol* 2007, 60:324–329
  259. KIVILUOTO T, SIREN J, LUUKKONEN P, KIVILAAKSO E. Randomised trial of laparoscopic versus open cholecystectomy for acute and gangrenous cholecystitis. *Lancet* 1998, 351:321–325
  260. HUPPERETS MD, VERHAGEN EA, VAN MECHELEN W. Effect of unsupervised home based proprioceptive training on recurrences of ankle sprain: Randomised controlled trial. *Br Med J* 2009, 339:b2684
  261. GARBER J, CLARKE GN, WEERSING VR, BEARDSLEE WR, BRENT DA, GLADSTONE TR ET AL. Prevention of depression in at-risk adolescents: A randomized controlled trial. *JAMA* 2009, 301:2215–2224
  262. CAMPBELL DT. Factors relevant to the validity of experiments in social settings. *Psychol Bull* 1957, 54:297–312
  263. ROTHWELL PM. Factors that can affect the external validity of randomised controlled trials. *PLoS Clin Trials* 2006, 1:e9
  264. KING M, NAZARETH I, LAMPE F, BOWER P, CHANDLER M, MOROU M ET AL. Conceptual framework and systematic review of the effects of participants' and professionals' preferences in randomised controlled trials. *Health Technol Assess* 2005, 9:1–186, iii–iv
  265. DJULBEGOVIC B, LACEVIC M, CANTOR A, FIELDS KK, BENNETT CL, ADAMS JR ET AL. The uncertainty principle and industry-sponsored research. *Lancet* 2000, 356:635–638
  266. DANS AL, DANS LF, GUYATT GH, RICHARDSON S. Users' guides to the medical literature: XIV. How to decide on the applicability of clinical trial results to your patient. Evidence-Based Medicine Working Group. *JAMA* 1998, 279:545–549
  267. SMITH GD, EGGER M. Who benefits from medical interventions? *Br Med J* 1994, 308:72–74
  268. McALISTER FA. Applying the results of systematic reviews at the bedside. In: Egger M, Davey Smith G, Altman DG (eds) *Systematic reviews in health care: Meta-analysis in context*. 2nd ed. BMJ Books, London, 2001
  269. BARTLETT C, DOYAL L, EBRAHIM S, DAVEY P, BACHMANN M, EGGER M ET AL. The causes and effects of socio-demographic exclusions from clinical trials. *Health Technol Assess* 2005, 9:iii–iv, ix–x, 1–152
  270. BONELL C, OAKLEY A, HARGREAVES J, STRANGE V, REES R. Assessment of generalisability in trials of health interventions: Suggested framework and systematic review. *Br Med J* 2006, 333:346–349
  271. BORNHÖFT G, MAXION-BERGEMANN S, WOLF U, KIENLE GS, MICHALSEN A, VOLLMAR HC ET AL. Checklist for the qualitative evaluation of clinical studies with particular focus on external validity and model validity. *BMC Med Res Methodol* 2006, 6:56
  272. LAUPACIS A, SACKETT DL, ROBERTS RS. An assessment of clinically useful measures of the consequences of treatment. *N Engl J Med* 1988, 318:1728–1733
  273. ALTMAN DG. Confidence intervals for the number needed to treat. *Br Med J* 1998, 317:1309–1312
  274. FANAROFF AA, KORONES SB, WRIGHT LL, WRIGHT EC, POLAND RL, BAUER CB ET AL. A controlled trial of intravenous immune globulin to reduce nosocomial infections in very-low-birth-weight infants. National Institute of Child Health and Human Development Neonatal Research Network. *N Engl J Med* 1994, 330:1107–1113
  275. ANONYMOUS. Randomised trial of intravenous atenolol among 16,027 cases of suspected acute myocardial infarction: ISIS-1. First International Study of Infarct Survival Collaborative Group. *Lancet* 1986, ii:57–66
  276. GÖTZSCHE PC, GJØRUP I, BONNÉN H, BRAHE NE, BECKER U, BURCHARTH F. Somatostatin vs placebo in bleeding oesophageal varices: Randomised trial and meta-analysis. *Br Med J* 1995, 310:1495–1498
  277. CLARKE M, HOPEWELL S, CHALMERS I. Reports of clinical trials should begin and end with up-to-date systematic reviews of other relevant evidence: A status report. *JR Soc Med* 2007, 100:187–190
  278. GOODMAN SN. Toward evidence-based medical statistics. 1: The P value fallacy. *Ann Intern Med* 1999, 130:995–1004
  279. GÖTZSCHE PC. Reference bias in reports of drug trials. *Br Med J (Clin Res Ed)* 1987, 295:654–656
  280. PAPP K, BISSONNETTE R, ROSOPH L, WASEL N, LYNDE CW, SEARLES G ET AL. Efficacy of ISA247 in plaque psoriasis: A randomised, multicentre, double-blind, placebo-controlled phase III study. *Lancet* 2008, 371:1337–1342
  281. DICKERSIN K. How important is publication bias? A synthesis of available data. *AIDS Educ Prev* 1997, 9(Suppl 1):15–21
  282. SONG F, EASTWOOD AJ, GILBODY S, DULEY L, SUTTON AJ. Publication and related biases. *Health Technol Assess* 2000, 4:1–115
  283. WILLIAMSON PR, GAMBLE C. Identification and impact of outcome selection bias in meta-analysis. *Stat Med* 2005, 24:1547–1561
  284. TRAMÈR MR, REYNOLDS DJ, MOORE RA, McQUAY HJ. Impact of covert duplicate publication on meta-analysis: A case study. *Br Med J* 1997, 315:635–640
  285. SIMES RJ. Publication bias: The case for an international registry of clinical trials. *J Clin Oncol* 1986, 4:1529–1541
  286. CHALMERS I. From optimism to disillusion about commitment to transparency in the medico-industrial complex. *JR Soc Med* 2006, 99:337–341
  287. TONKS A. A clinical trials register for Europe. *Br Med J* 2002, 325:1314–1315
  288. DICKERSIN K, RENNIE D. Registering clinical trials. *JAMA* 2003, 290:516–523
  289. WHITTINGTON CJ, KENDALL T, FONAGY P, COTTRELL D, COTGROVE A, BODDINGTON E. Selective serotonin reuptake inhibitors in childhood depression: Systematic review of published versus unpublished data. *Lancet* 2004, 363:1341–1345
  290. DE ANGELIS CD, DRAZEN JM, FRIZELLE FA, HAUG C, HOEY J, HORTON R ET AL. Is this clinical trial fully registered? A statement from the International Committee of Medical Journal Edi-



- tors. *Lancet* 2005, 365:1827–1829
291. ZARIN DA, IDE NC, TSET, HARLAN WR, WEST JC, LINDBERG DA. Issues in the registration of clinical trials. *JAMA* 2007, 297:2112–2120
292. HOPEWELL S, ALTMAN DG, MOHER D, SCHULZ KF. Endorsement of the CONSORT statement by high impact factor medical journals: A survey of journal editors and journal “instructions to authors”. *Trials* 2008, 9:20
293. RUSSELL JA, WALLEY KR, SINGER J, GORDON AC, HEBERT PC, COOPER DJ ET AL. Vasopressin versus norepinephrine infusion in patients with septic shock. *N Engl J Med* 2008, 358:877–887
294. CHAN AW, TETZLAFF J, ALTMAN D, GÖTZSCHE PC, HRÓBJARTSSON A, KRLEZA-JERIC K ET AL. The SPIRIT initiative: Defining standard protocol items for randomised trials. Oral presentation at the 16th Cochrane Colloquium: Evidence in the era of globalisation; 2008 Oct 3–7; Freiburg, Germany [abstract]. *Z Evid Fortbild Qual Gesundheitswes* 2008, 102:27
295. GREGSON S, ADAMSON S, PAPAYA S, MUNDONDO J, NYAMUKA-PA CA, MASON PR ET AL. Impact and process evaluation of integrated community and clinic-based HIV-1 control: A cluster-randomised trial in eastern Zimbabwe. *PLoS Med* 2007, 4:e102
296. SHADDY RE, BOUCEK MM, HSU DT, BOUCEK RJ, CANTER CE, MAHONY L ET AL. Carvedilol for children and adolescents with heart failure: A randomized controlled trial. *JAMA* 2007, 298:1171–1179
297. LEXCHIN J, BERO LA, DJULBEGOVIC B, CLARK O. Pharmaceutical industry sponsorship and research outcome and quality: Systematic review. *Br Med J* 2003, 326:1167–1170
298. KJAERGARD LL, ALS-NIELSEN B. Association between competing interests and authors’ conclusions: Epidemiological study of randomised clinical trials published in the BMJ. *Br Med J* 2002, 325:249
299. BERO L, OOSTVOGEL F, BACCHETTI P, LEE K. Factors associated with findings of published trials of drug-drug comparisons: Why some statins appear more efficacious than others. *PLoS Med* 2007, 4:e184
300. SISMONDO S. Pharmaceutical company funding and its consequences: A qualitative systematic review. *Contemp Clin Trials* 2008, 29:109–113
301. ALS-NIELSEN B, CHEN W, GLUUD C, KJAERGARD LL. Association of funding and conclusions in randomized drug trials: A reflection of treatment effect or adverse events? *JAMA* 2003, 290:921–928
302. ROSS JS, HILL KP, EGILMAN DS, KRUMHOLZ HM. Guest authorship and ghost writing in publications related to rofecoxib: A case study of industry documents from rofecoxib litigation. *JAMA* 2008, 299:1800–1812
303. McALISTER FA, STRAUS SE, SACKETT DL, ALTMAN DG. Analysis and reporting of factorial trials: A systematic review. *JAMA* 2003, 289:2545–2553
304. SENN S. *Cross-over trials in clinical research*. 2nd ed. Wiley, New York, 2002
305. DEEKS JJ, DINNES J, D’AMICO R, SOWDEN AJ, SAKAROVITCH C, SONG F ET AL. Evaluating non-randomised intervention studies. *Health Technol Assess* 2003, 7:iii–x, 1–173
306. KUNZ R, VIST G, OXMAN AD. Randomisation to protect against selection bias in healthcare trials. *Cochrane Database Syst Rev* 2007, 2:MR000012
307. COLLINS R, MACMAHON S. Reliable assessment of the effects of treatment on mortality and major morbidity. I: Clinical trials. *Lancet* 2001, 357:373–380
308. SCHULZ KF. Randomised trials, human nature, and reporting guidelines. *Lancet* 1996, 348:596–598
309. MURRAY GD. Promoting good research practice. *Stat Methods Med Res* 2000, 9:17–24
310. NARAHARI SR, RYAN TJ, AGGITHAYA MG, BOSE KS, PRASANNA KS. Evidence-based approaches for the Ayurvedic traditional herbal formulations: Toward an Ayurvedic CONSORT model. *J Altern Complement Med* 2008, 14:769–776
311. BOSSUYT PM, REITSMA JB, BRUNS DE, GATSONIS CA, GLASZIOU PP, IRWIG LM ET AL. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD initiative. *Ann Intern Med* 2003, 138:40–44
312. VON ELM E, ALTMAN DG, EGGER M, POCOCK SJ, GÖTZSCHE PC, VANDENBROUCKE JP. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Ann Intern Med* 2007, 147:573–577
313. TONINO PA, DE BRUYNE B, PIJLS NH, SIEBERT U, IKENO F, VAN’T VEER M ET AL. Fractional flow reserve versus angiography for guiding percutaneous coronary intervention. *N Engl J Med* 2009, 360:213–224
- Corresponding author:*
- A.B. Haidich, Department of Hygiene and Epidemiology, Aristotle University of Thessaloniki, School of Medicine, GR-541 24 Thessaloniki, Greece  
e-mail: haidich@med.auth.gr

## Πλαίσιο 1 | Διανομή θεραπείας. Τι είναι τόσο ιδιαίτερο σχετικά με την τυχαιοποίηση;

Η μέθοδος που χρησιμοποιείται για να ταξινομηθούν οι παρεμβάσεις στους συμμετέχοντες της δοκιμής είναι μια κρίσιμη πτυχή του σχεδιασμού μιας κλινικής δοκιμής. Η τυχαία ταξινόμηση είναι η προτιμώμενη μέθοδος. Έχει χρησιμοποιηθεί επιτυχώς τακτικά στις δοκιμές για >50 έτη.<sup>24</sup> Η τυχαιοποίηση έχει τρία σημαντικά πλεονεκτήματα:<sup>25</sup> Πρώτον, όταν εφαρμόζεται κατάλληλα, εξαλείφει το σφάλμα επιλογής, ισορροπώντας γνωστούς και άγνωστους προγνωστικούς παράγοντες, στην ταξινόμηση των θεραπειών. Χωρίς τυχαιοποίηση, οι συγκρίσεις θεραπειών μπορεί να είναι προκατειλημμένες, είτε συνειδητά είτε όχι, από την επιλογή των συμμετεχόντων ενός συγκεκριμένου είδους για να λάβει μια συγκεκριμένη θεραπεία. Δεύτερον, η τυχαία ταξινόμηση επιτρέπει τη χρήση της θεωρίας των πιθανοτήτων για να εκφράσει την πιθανότητα ότι οποιαδήποτε διαφορά στην έκβαση μεταξύ των ομάδων παρέμβασης απεικονίζει μερικώς τύχη.<sup>26</sup> Τρίτον, η τυχαία διανομή, σε μερικές καταστάσεις, διευκολύνει την τυφλοποίηση της ταυτότητας των θεραπειών στους ερευνητές, στους συμμετέχοντες και στους αξιολογητές, ενδεχομένως με τη χρήση ενός εικονικού φαρμάκου, που μειώνει το σφάλμα μετά από την ταξινόμηση των θεραπειών.<sup>27</sup> Από αυτά τα τρία πλεονεκτήματα, η μείωση του σφάλματος επιλογής κατά την είσοδο στη δοκιμή είναι συνήθως το περισσότερο σημαντικό.<sup>28</sup>

Η επιτυχής τυχαιοποίηση στην πράξη εξαρτάται από δύο συσχετισμένες πτυχές, επαρκή παραγωγή μιας μη προβλέψιμης ακολουθίας διανομής και απόκρυψη αυτής της ακολουθίας μέχρι να γίνει η ταξινόμηση.<sup>2,23</sup> Ένα θεμάκλειδί είναι εάν το πρόγραμμα είναι γνωστό ή προβλέψιμο από τους ανθρώπους που εμπλέκονται στη διανομή των συμμετεχόντων σε ομάδες σύγκρισης.<sup>29</sup> Το σύστημα διανομής της θεραπείας πρέπει να οργανωθεί έτσι ώστε το άτομο που εγγράφει τους συμμετέχοντες να μη γνωρίζει εκ των προτέρων ποια θεραπεία θα λάβει το επόμενο πρόσωπο, μια διαδικασία που καλείται απόκρυψη διανομής.<sup>2,23</sup> Η απόκρυψη κατάλληλης διανομής προστατεύει τη γνώση των επικείμενων ταξινόμησεων, δεδομένου ότι οι κατάλληλες τυχαιές ακολουθίες αποτρέπουν τη σωστή πρόβλεψη των μελλοντικών ταξινόμησεων βασισμένων στη γνώση των προηγούμενων ταξινόμησεων.

## Πλαίσιο 2 | Τυχαιοποίηση και ελαχιστοποίηση

*Απλή τυχαιοποίηση:* Η καθαρή τυχαιοποίηση βασισμένη σε μια μοναδική αναλογία διανομής είναι γνωστή ως απλή τυχαιοποίηση. Η απλή τυχαιοποίηση με μια 1:1 αναλογία διανομής είναι ανάλογη με μια ρίψη νομίσματος, αν και δεν

υποστηρίζουμε τη ρίψη νομίσματος για την τυχαιοποίηση σε μια ΤΚΔ. Ο όρος «απλή» είναι κατά κάποιον τρόπο μια ακυριολεξία. Ενώ άλλα σχήματα τυχαιοποίησης ακούγονται σύνθετα και πιο εξεζητημένα, στην πραγματικότητα, η απλή τυχαιοποίηση είναι κομψά εξεζητημένη, δεδομένου ότι είναι πιο απρόβλεπτη και υπερβαίνει τα επίπεδα πρόληψης του σφάλματος όλων των άλλων εναλλακτικών τυχαιοποιήσεων.

*Περιορισμένη τυχαιοποίηση:* Οποιαδήποτε τυχαιοποιημένη προσέγγιση που δεν είναι απλή τυχαιοποίηση. Η τυχαιοποίηση με μπλοκ είναι η πιο κοινή μορφή. Άλλες έννοιες της περιορισμένης τυχαιοποίησης περιλαμβάνουν την αντικατάσταση, το μεροληπτικό κέρμα, και την τυχαιοποίηση δοχείων, αν και αυτοί χρησιμοποιούνται πολύ λιγότερο συχνά.<sup>141</sup>

*Τυχαιοποίηση με μπλοκ:* Χρησιμοποιείται για να εξασφαλίσει ότι οι ομάδες σύγκρισης θα παραχθούν σύμφωνα με μια προκαθορισμένη αναλογία, συνήθως 1:1 ή ομάδες περίπου του ίδιου μεγέθους. Η τυχαιοποίηση με μπλοκ μπορεί να χρησιμοποιηθεί για να εξασφαλίσει στενή ισορροπία των αριθμών σε κάθε ομάδα, οποιαδήποτε στιγμή κατά τη διάρκεια της δοκιμής. Για κάθε μπλοκ 8 συμμετεχόντων, για παράδειγμα, 4 θα διανέμονται σε κάθε σκέλος της δοκιμής.<sup>142</sup> Η βελτιωμένη ισορροπία έρχεται με κόστος τη μείωση της μη προβλεψιμότητας της ακολουθίας. Αν και η σειρά των παρεμβάσεων ποικίλλει τυχαιοποιημένα μέσα σε κάθε μπλοκ, ένα πρόσωπο που τρέχει τη δοκιμή θα μπορούσε να συναγάγει μερικές από τις επόμενες διανομές θεραπειών, εάν γνώριζε το μέγεθος του μπλοκ.<sup>143</sup> Η τυφλοποίηση των παρεμβάσεων, η χρήση μπλοκ μεγαλύτερων μεγεθών και η τυχαία διαφοροποίηση του μεγέθους των μπλοκ μπορούν να βελτιώσουν αυτό το πρόβλημα.

*Διαστρωματοποιημένη τυχαιοποίηση:* Η διαστρωμάτωση χρησιμοποιείται για να εξασφαλίσει καλή ισορροπία των χαρακτηριστικών των συμμετεχόντων σε κάθε ομάδα. Κατά τύχη, ιδιαίτερα στις μικρές δοκιμές, οι ομάδες της μελέτης μπορεί να μην είναι καλά αντιστοιχισμένες για τα βασικά χαρακτηριστικά, όπως η ηλικία και το στάδιο της νόσου. Αυτό αποδυναμώνει την εγκυρότητα της δοκιμής.<sup>144</sup> Τέτοιες δυσαναλογίες μπορούν να αποφευχθούν χωρίς θυσία των πλεονεκτημάτων της τυχαιοποίησης. Η διαστρωμάτωση εξασφαλίζει ότι οι αριθμοί των συμμετεχόντων που λαμβάνουν κάθε παρέμβαση είναι στενά ισορροπημένοι μέσα σε κάθε στρώμα. Η διαστρωματοποιημένη τυχαιοποίηση επιτυγχάνεται με την εκτέλεση μιας χωριστής διαδικασίας τυχαιοποίησης μέσα σε κάθε ένα από δύο ή περισσότερα υποσύνολα συμμετεχόντων (για παράδειγμα, εκείνων που καθορίζουν κάθε κέντρο μελέτης, την ηλικία ή τη

σοβαρότητα των νόσων). Η διαστρωμάτωση ανά κέντρο είναι κοινή στις πολυκεντρικές δοκιμές. Η διαστρωμάτωση απαιτεί κάποια μορφή περιορισμού (όπως την τυχαιοποίηση με μπλοκ μέσα στα στρώματα). Η διαστρωμάτωση χωρίς τυχαιοποίηση με μπλοκ δεν είναι αποτελεσματική.

**Ελαχιστοποίηση:** Η ελαχιστοποίηση εξασφαλίζει ισορροπία μεταξύ των ομάδων παρέμβασης για ορισμένους επιλεγμένους παράγοντες ασθενών (όπως η ηλικία).<sup>22,60</sup> Ο πρώτος ασθενής είναι πραγματικά διανεμημένος τυχαιοποιημένα. Για κάθε επόμενο συμμετέχοντα, η διανομή θεραπείας που ελαχιστοποιεί την ανισορροπία στους επιλεγμένους παράγοντες μεταξύ των ομάδων αναγνωρίζεται εκείνη τη στιγμή. Η διανομή αυτή μπορεί έπειτα να χρησιμοποιηθεί ή μια επιλογή μπορεί να γίνει τυχαία με μια μεγάλη στάθμιση υπέρ της παρέμβασης που θα ελαχιστοποιούσε την ανισορροπία (για παράδειγμα, με μια πιθανότητα 0,8). Η χρήση μιας τυχαίας σύνθεσης είναι γενικά προτιμητέα. Η ελαχιστοποίηση έχει το πλεονέκτημα της δημιουργίας μικρών ομάδων, πολύ παρόμοιων από την άποψη των χαρακτηριστικών των συμμετεχόντων σε όλα τα στάδια της δοκιμής. Η ελαχιστοποίηση προσφέρει τη μόνη αποδεκτή εναλλακτική λύση στην τυχαιοποίηση, ενώ μερικοί έχουν υποστηρίξει ότι είναι ανώτερη.<sup>145</sup> Από την άλλη πλευρά, η ελαχιστοποίηση στερείται της θεωρητικής βάσης για την εξάλειψη του σφάλματος σε όλους τους γνωστούς και άγνωστους παράγοντες. Εν τούτοις, γενικά, οι δοκιμές που χρησιμοποιούν ελαχιστοποίηση θεωρούνται μεθοδολογικά ισοδύναμες με τις τυχαιοποιημένες δοκιμές, ακόμη και όταν δεν ενσωματώνεται ένα τυχαίο στοιχείο.

### Πλαίσιο 3 | Βήματα σε μια τυπική διαδικασία τυχαιοποίησης

#### Παραγωγή ακολουθίας

- Παράγετε την ακολουθία κατανομής με κάποια τυχαία διαδικασία.

#### Απόκρυψη διανομής

- Αναπτύξτε το μηχανισμό απόκρυψης διανομής (όπως τα αριθμημένα, πανομοιότυπα φιαλίδια ή οι διαδοχικά αριθμημένοι, σφραγισμένοι, αδιαφανείς φάκελοι)
- Προετοιμάστε το μηχανισμό απόκρυψης διανομής, χρησιμοποιώντας την ακολουθία διανομής από το βήμα της παραγωγής ακολουθίας.

#### Εφαρμογή

- Εγγράψτε τους συμμετέχοντες:

Αξιολογήστε την καταλληλότητα

Συζητήστε τη δοκιμή

Λάβετε ενημερωμένη συγκατάθεση

Εγγράψτε το συμμετέχοντα στη δοκιμή

- Εξακριβώστε την ταξινόμηση παρέμβασης (όπως το άνοιγμα του επόμενου φακέλου)
- Χορηγήστε την παρέμβαση.

### Πλαίσιο 4 | Ορολογία τυφλοποίησης

Ένας τεχνικός όρος, για να έχει χρησιμότητα, πρέπει να έχει συνέπεια στη χρήση και στην ερμηνεία του. Οι συγγραφείς των δοκιμών χρησιμοποιούν συνήθως τον όρο «διπλή-τυφλή» και, λιγότερο συχνά, τους όρους «απλή-τυφλή» ή «τριπλή-τυφλή». Ένα πρόβλημα με αυτό το λεξιλόγιο είναι ότι υπάρχει μεγάλη μεταβλητότητα στις ερμηνείες κλινικών ιατρών και στους ορισμούς επιδημιολογικών εγχειριδίων αυτών των όρων.<sup>169</sup> Επί πλέον, σε μια μελέτη 200 ΤΚΔ που αναφέρονται ως διπλές-τυφλές βρέθηκαν 18 διαφορετικοί συνδυασμοί ομάδων, πραγματικά τυφλοποιημένων όταν ερωτήθηκαν οι συγγραφείς αυτών των δοκιμών, και περίπου μία σε κάθε πέντε από αυτές τις δοκιμές –που αναφέρθηκαν ως διπλές-τυφλές– δεν τυφλοποίησε τους συμμετέχοντες, τους προμηθευτές υπηρεσιών υγείας ή τους συλλέκτες δεδομένων.<sup>170</sup>

Η συγκεκριμένη έρευνα δείχνει ότι οι όροι είναι διφορούμενοι και, έτσι, οι συγγραφείς και οι εκδότες πρέπει να εγκαταλείψουν τη χρήση τους. Οι συγγραφείς πρέπει αντ' αυτού να αναφέρουν ρητά την κατάσταση τυφλοποίησης των ατόμων που εμπλέκονται, για τους οποίους η τυφλοποίηση μπορεί να επηρεάσει την αξιοπιστία μιας δοκιμής.

Οι προμηθευτές υπηρεσιών υγείας περιλαμβάνουν όλο το προσωπικό (για παράδειγμα, ιατροί, χειροπράκτες, φυσικοθεραπευτές, νοσηλεύτριες) που φροντίζουν τους συμμετέχοντες κατά τη διάρκεια της δοκιμής. Οι συλλέκτες δεδομένων είναι τα άτομα που συλλέγουν τα δεδομένα για τις εκβάσεις των δοκιμών. Οι κριτές εκβάσεων είναι τα άτομα που καθορίζουν εάν σε ένα συμμετέχοντα δεν συνέβησαν οι εκβάσεις που μας ενδιαφέρουν.

Μερικοί ερευνητές έχουν υποστηρίξει επίσης την τυφλοποίηση και την καταγραφή της κατάστασης της τυφλοποίησης της επιτροπής ελέγχου δεδομένων και των συγγραφέων της εργασίας.<sup>160</sup> Η τυφλοποίηση αυτών των ομάδων είναι ασυνήθης και η αξία της τυφλοποίησής τους τελεί υπό συζήτηση.<sup>171</sup>

Μερικές φορές, μια ομάδα ατόμων (όπως οι προμηθευτές υπηρεσιών υγείας) είναι τα ίδια άτομα που εκπληρώνουν έναν άλλο ρόλο σε μια δοκιμή (όπως οι συλλέκτες δεδο-

μένων). Ακόμη και σε αυτή την περίπτωση, οι συγγραφείς πρέπει να δηλώνουν ρητά την κατάσταση τυφλοποίησης αυτών των ομάδων για να επιτρέψουν στους αναγνώστες να κρίνουν την εγκυρότητα της δοκιμής.

## Πλαίσιο 5 | Πρόωρη διακοπή

Οι ΤΚΔ μπορεί να σταματήσουν, όταν φθάσουν στο προοριζόμενο μέγεθος του δείγματος, στον προοριζόμενο αριθμό συμβαμάτων, στην προοριζόμενη διάρκεια παρακολούθησης ή όταν φθάσουν στην προγραμματισμένη ημερομηνία τερματισμού. Εναλλακτικά, οι ΤΚΔ ενδέχεται να διακοπούν νωρίτερα από το προγραμματισμένο, επειδή το αποτέλεσμα της ενδιάμεσης ανάλυσης έδειξε μεγαλύτερο από το αναμενόμενο όφελος ή βλαβερές συνέπειες στην πειραματική παρέμβαση. Επίσης, οι ΤΚΔ μπορεί να διακοπούν νωρίτερα, αν οι ερευνητές εντοπίσουν ενδείξεις μη σημαντικής διαφοράς ανάμεσα στην πειραματική παρέμβαση και την παρέμβαση ελέγχου (αυτό σημαίνει διακοπή λόγω ματαιότητας). Συμπληρωματικά, οι δοκιμές μπορεί να διακοπούν πρώιμα, επειδή η δοκιμή καθίσταται μη βιώσιμη: Η χρηματοδότηση εξαφανίζεται, οι ερευνητές δεν μπορούν να βρουν κατάλληλους ασθενείς ή οι παρεμβάσεις της μελέτης ή τα αποτελέσματα άλλων μελετών καταστήσουν το ερευνητικό ερώτημα μη σχετικό.

Πλήρης αναφορά του λόγου διακοπής της δοκιμής είναι σημαντική για λήψη απόφασης βασισμένης σε ενδείξεις (βλέπε στοιχείο 14β). Οι ερευνητές που εξέτασαν γιατί διακόπηκαν 143 δοκιμές νωρίς για όφελος έδειξαν ότι πολλές δεν ανέφεραν καίριες μεθοδολογικές πληροφορίες σχετικά με το πώς κατέληξαν στην απόφαση διακοπής το σχεδιαζόμενο μέγεθος δείγματος ( $n=28$ ), την ενδιάμεση ανάλυση μετά την οποία η δοκιμή διακόπηκε ( $n=45$ ) ή αν ένας κανόνας διακοπής υποστήριζε την απόφαση ( $n=48$ ).<sup>134</sup> Το στοιχείο 7β του καταλόγου απαιτεί την αναφορά της χρονικής στιγμής των ενδιάμεσων αναλύσεων, τι τις πυροδότησε, πόσες έλαβαν χώρα, εάν είχαν σχεδιαστεί ή ήταν εκ των υστέρων (*ad hoc*), και αν υπήρχαν εκ των προτέρων (*a priori*) στατιστικές κατευθυντήριες οδηγίες και κανόνες διακοπής. Επί πλέον, είναι χρήσιμο να γνωρίζει κάποιος εάν συμμετείχε στην ανάλυση μια ανεξάρτητη επιτροπή παρακολούθησης των δεδομένων (και από ποιους συγκροτήθηκε, με ιδιαίτερη προσοχή στο ρόλο της πηγής χρηματοδότησης) και ποιος έλαβε την απόφαση για διακοπή. Συχνά, η επιτροπή ασφάλειας και επιτήρησης δεδομένων προτείνει και οι χρηματοδότες (χορηγοί) ή οι ερευνητές να λαμβάνουν την απόφαση για διακοπή.

Οι δοκιμές οι οποίες διακόπτονται πρώιμα για λόγους εμφανώς ανεξάρτητους από τα ευρήματα της δοκιμής και οι δοκιμές που φθάνουν στο σχεδιασμένο τερματισμό, είναι

απίθανο να εμφανίσουν σφάλματα από τη διακοπή.<sup>207</sup> Σε αυτές τις περιπτώσεις, οι συγγραφείς πρέπει να αναφέρουν αν έγινε ενδιάμεση ανάλυση και αν τα αποτελέσματα αυτά ήταν διαθέσιμα στο χρηματοδότη.

Η πίεση για δοκιμές που αλλάζουν την παρέμβαση ανάλογα με την ενδιάμεση ανάλυση, κι έτσι τη δυνατότητα ταχύτερης αξιολόγησης μιας υποσχόμενης παρέμβασης για ταχέως εξελισσόμενες και μοιραίες καταστάσεις, απαιτούν ακόμη πιο προσεκτική αναφορά της διαδικασίας και της απόφασης για πρόωρη διακοπή της δοκιμής.<sup>208</sup>

## Πλαίσιο 6 | Ανάλυση πρόθεσης για θεραπεία

Η ιδιαίτερη ισχύς των ΤΚΔ είναι η αποφυγή σφαλμάτων, όταν διανέμονται οι παρεμβάσεις στους συμμετέχοντες στη δοκιμή (βλέπε πλαίσιο 1). Αυτή η ισχύς επιτρέπει σημαντικά συμπεράσματα σχετικά με το αίτιο και το αποτέλεσμα, τα οποία δεν δικαιολογούνται σε άλλο σχεδιασμό δοκιμής. Προκειμένου να διατηρηθεί ακέραιο το τεράστιο όφελος της τυχαιοποίησης, στην ανάλυση πρέπει να συμπεριληφθούν όλοι οι τυχαιοποιημένοι συμμετέχοντες, διατηρώντας τους όλους στην ομάδα στην οποία είχαν διανεμηθεί. Αυτοί οι δύο παράγοντες καθορίζουν την ανάλυση πρόθεσης για θεραπεία, η οποία προτείνεται ευρέως ως η επιθυμητή στρατηγική ανάλυσης.<sup>18,223</sup> Η ανάλυση πρόθεσης για θεραπεία ανταποκρίνεται στην ανάλυση των ομάδων, ακριβώς όπως έχουν τυχαιοποιηθεί. Αυστηρή ανάλυση πρόθεσης για θεραπεία είναι δύσκολο να επιτευχθεί για δύο κύριους λόγους: Ελλιπείς εκβάσεις για κάποιους από τους συμμετέχοντες και μη προσκόλληση στο πρωτόκολλο της δοκιμής.

### Ελλιπείς εκβάσεις

Πολλοί από τους διεξάγοντες τις δοκιμές αποκλείουν τους ασθενείς χωρίς μια παρατηρούμενη έκβαση. Συχνά, αυτό είναι λογικό, αλλά τη στιγμή που κάποιοι τυχαιοποιημένοι ασθενείς αποκλείονται, η ανάλυση παύει να είναι αυστηρά ανάλυση πρόθεσης για θεραπεία. Πράγματι, οι περισσότερες τυχαιοποιημένες μελέτες έχουν μερικές ελλιπείς παρατηρήσεις. Οι διεξάγοντες τις δοκιμές αποτελεσματικά πρέπει να επιλέξουν μεταξύ της παράλειψης των συμμετεχόντων χωρίς δεδομένα τελικής έκβασης ή της αντικατάστασης των ελλιπών δεδομένων έκβασης τους.<sup>225</sup> Μια «ολοκληρωμένη ανάλυση» (ή «διαθέσιμη ανάλυση») περιλαμβάνει μόνο τα δεδομένα των οποίων οι εκβάσεις είναι γνωστές. Μολονότι μερικές ελλιπείς εκβάσεις δεν δημιουργούν πρόβλημα, στις μισές από τις δοκιμές ποσοστό >10% των τυχαιοποιημένων ασθενών μπορεί να έχουν ελλιπείς εκβάσεις.<sup>226</sup> Αυτή η συνήθης προσέγγιση χάνει σε ισχύ μειώνοντας το μέγεθος του δείγματος, ενώ



μπορεί να υπεισέλθουν σφάλματα, εάν το να χαθεί κάποιος από την παρακολούθηση σχετίζεται με την ανταπόκριση του ασθενούς στη θεραπεία. Υπάρχει ανησυχία όταν η συχνότητα ή οι αιτίες της απώλειας διαφέρουν ανάμεσα στις ομάδες της παρέμβασης.

Συμμετέχοντες με ελλειπίς εκβάσεις μπορεί να περιληφθούν στην ανάλυση μόνο αν οι εκβάσεις τους έχουν αντικατασταθεί (imputed) (αυτό σημαίνει ότι οι εκβάσεις υπολογίζονται από άλλες πληροφορίες που έχουν συλλεχθεί). Η αντικατάσταση των ελλειπών δεδομένων (imputation) επιτρέπει στην ανάλυση να συμμορφώνεται με την ανάλυση πρόθεσης για θεραπεία, αλλά απαιτεί ισχυρές παραδοχές, οι οποίες είναι δύσκολο να δικαιολογηθούν.<sup>227</sup> Απλές τεχνικές αντικατάστασης είναι ελκυστικές, αλλά δεν συστήνεται η χρήση τους. Συγκεκριμένα, μια ευρέως χρησιμοποιούμενη μέθοδος είναι «η τελευταία παρατήρηση να προωθηθεί (last observation carry forward)», κατά την οποία ελλειπίς τελικές τιμές της μεταβλητής έκβασης αντικαθίστανται από την τελευταία γνωστή τιμή, προτού χαθεί ο συμμετέχων από την παρακολούθηση. Αυτό είναι ελκυστικό εξ αιτίας της απλότητάς του, αλλά η μέθοδος ενδέχεται να εισάγει σφάλματα<sup>228</sup> και καμιά παραδοχή δεν γίνεται για την αβεβαιότητα της αντικατάστασης.<sup>229</sup> Πολλοί συγγραφείς έχουν ασκήσει αυστηρή κριτική στη μέθοδο της προώθησης της τελευταίας παρατήρησης.<sup>229–231</sup>

### Μη προσκόλληση στο πρωτόκολλο

Ένα ξεχωριστό θέμα είναι ότι το πρωτόκολλο της δοκιμής ενδέχεται να μην έχει ακολουθηθεί πλήρως, για κάποιους από τους συμμετέχοντες στη δοκιμή. Σύνηθες παράδειγμα αποτελεί το γεγονός ότι οι συμμετέχοντες που

δεν πληρούσαν τα κριτήρια εισόδου (όπως η λανθασμένη διάγνωση, πολύ νέοι) έλαβαν μια απαγορευμένη συν-παρέμβαση, δεν έλαβαν όλη την προορισμένη θεραπεία ή έλαβαν μια διαφορετική θεραπεία ή καμιά παρέμβαση. Ένας απλός τρόπος χειρισμού κάθε παράκαμψης από το πρωτόκολλο είναι να αγνοηθούν: Όλοι οι συμμετέχοντες μπορούν να συμπεριληφθούν στην ανάλυση, ανεξάρτητα από την προσκόλληση στο πρωτόκολλο, και αυτή είναι η προσέγγιση πρόθεσης για θεραπεία. Έτσι, ο αποκλεισμός οποιουδήποτε συμμετέχοντα για τέτοιο λόγο είναι ασύμβατος με την ανάλυση πρόθεσης για θεραπεία.

Ο όρος «τροποποιημένη πρόθεση για θεραπεία» είναι σχετικά ευρέως χρησιμοποιούμενος για την περιγραφή μιας ανάλυσης η οποία αποκλείει συμμετέχοντες που δεν έχουν επαρκή προσκόλληση στο πρωτόκολλο, ιδιαίτερα αυτούς οι οποίοι δεν έλαβαν μια καθορισμένη ελάχιστη παρέμβαση.<sup>232</sup> Ένας εναλλακτικός όρος είναι «ανάλυση σύμφωνα με το πρωτόκολλο (per protocol)». Παρ' ότι μια ανάλυση σύμφωνα με το πρωτόκολλο είναι κατάλληλη για κάποιες συνθήκες, πρέπει να φέρει τη σωστή επισήμανση, ως μη τυχαιοποιημένη σύγκριση παρατήρησης. Κάθε αποκλεισμός συμμετεχόντων από την ανάλυση θέτει συμβιβασμούς στην τυχαιοποίηση και πιθανόν να οδηγήσει σε σφάλματα στα αποτελέσματα.

Όπως «η πρόθεση για θεραπεία», καμιά από τις άλλες επισημάνσεις δεν αποσαφηνίζει αξιόπιστα ποιοι ασθενείς συμπεριελήφθησαν. Έτσι, στον κατάλογο του CONSORT έχουν αφαιρεθεί οι συγκεκριμένες απαιτήσεις για ανάλυση πρόθεσης για θεραπεία προς όφελος μιας σαφούς περιγραφής των ατόμων που ακριβώς συμπεριελήφθησαν στην κάθε ανάλυση.

.....